

Hollow arguments against nuclear test ban

The Senate's approach to the Comprehensive Test Ban Treaty will show whether the United States is ready to lead a world emerging from the deceptive simplicities of the Cold War.

The end of nuclear testing in 1992 was, for many, a powerful sign that a more enlightened world might emerge following the collapse of the Soviet Union. The Comprehensive Test Ban Treaty (CTBT), which was signed by 44 countries in September last year, is intended to secure the current voluntary moratorium and must now be ratified by the signatories' individual legislatures in order to take effect. In the United States, President Bill Clinton has just sent it to the Senate, where the support of 67 senators is needed to secure ratification. Without US ratification, the treaty will fail. But as preliminary hearings last week made clear, those 67 votes may prove elusive.

The Clinton administration has laid careful groundwork to build support for ratification. It has pledged to spend \$4.5 billion a year for ten years on an elaborate stockpile stewardship programme, in which the nuclear weapons laboratories will ensure the safety and reliability of existing weapons without testing. That programme was designed to persuade the laboratories to support the treaty and a cessation of tests, but it was never likely to assuage the most ardent proponents of testing. Relatively silent of late, they emerged last week, at the first Senate hearings in the run-up to treaty ratification. They include some nuclear weapons scientists — free to say whatever they like about the plans of their political masters — as well as assorted unreformed cold warriors.

Credibility

To those who would sleep easier at night in a world without nuclear weapons tests, the case in their favour can seem somewhat obscure. But there are at least three reasons which the treaty's opponents give for abandoning it and resuming the tests. Testing, they contend, is the only way adequately to ensure the safety and reliability of existing US weapons. The credibility of the stockpile, in the eyes of potential adversaries, is enhanced by regular testing, they say. The opponents also argue that the United States is technically incapable of verifying others' compliance with the treaty.

In truth, the safety of the US weapons stockpile is not at issue. It sounds good to couple 'safety and reliability' as a justification for testing or for the expensive stockpile stewardship programme. But there is no instance in America's lengthy experience of a safety problem related to ageing. Nor would any such problem be resolved by testing. Reliability, on the other hand, is indeed the main objective of the stockpile stewardship programme, and the main concern of its critics.

The directors of the weapons laboratories say the stewardship programme will ensure reliability, whereas treaty opponents say that testing would do it better. James Schlesinger, the former defence secretary, told the Senate last week that confidence in reliability will decrease little by little, so as eventually to erode the credibility of the weapons, but the United States has not used routine tests to ensure reliability in the past. Schlesinger also suggested that Saddam Hussein was deterred from using chemical weapons in the Gulf War by US readiness to use nuclear weapons in retaliation. It is perhaps a measure of the world-view of treaty opponents that they can make such unwarranted claims for the military worth of nuclear weapons. Perhaps what they really desire is

regular testing as a means to proclaim US military superiority — precisely the type of activity that should have ended with the Cold War.

Verification

The final argument being brought to bear against the CTBT is that the United States will be unable to verify the compliance of other countries. Treaty opponents have made this case, characteristically, by tossing a smoke-bomb into the debate, in the shape of their recent contention that an event near Novaya Zemlya, in northern Russia, was a surreptitious nuclear test.

Premature intelligence reports that Russia had carried out such a test on 16 August were leaked to the press by unnamed government officials. Subsequent study by lay seismologists showed conclusively that the event in question was an offshore earthquake (*Nature* 389, 425; 1997). Although several senators have used the alleged 'ambiguity' of this event to cast doubt on the ability of seismologists to track nuclear tests, the episode in fact serves to demonstrate the opposite.

The aftermath of the Novaya Zemlya incident highlights a thorny issue related to CTBT ratification in the United States — the question of whether verification will be conducted in secret by the Department of Defense and the Central Intelligence Agency (CIA), or in public, with the involvement of a wider scientific community. Although scientists would prefer as open a process as possible, steps may have to be taken to ensure that officials at the CIA and Department of Defense are also satisfied with the process. Responsibility for verification is just one of a number of issues on which the Senate may seek concessions from the administration in exchange for ratification.

Last week's hearings by two Senate subcommittees were but early skirmishes in what is likely to be a protracted ratification debate. The National Security and Intelligence committees will scrutinize the issue before the Foreign Relations Committee, chaired by Jesse Helms (Republican, North Carolina), decides whether to bring the treaty to the Senate floor. Helms surprised everyone last year by ignoring conservative opposition and lending his support to the Chemical Weapons Convention (CWC), which was duly ratified. In this Republican-controlled Senate, however, the CWC benefited from its origins with the administration of President George Bush. The CTBT is Clinton's animal, and prospects for its ratification remain highly uncertain.

Countries such as India are entitled to have reservations about a treaty that leaves the existing nuclear powers with large weapons stockpiles, and the United States, in particular, with an unchallengeable lead in nuclear weapons technology, based on computer simulation. Nevertheless, the arguments in favour of ratification by all the nuclear weapons states are overwhelming. The Senate's actions on this issue will critically define America's role in the 'new world order'. Outright rejection of the treaty risks a resumption of nuclear testing in the United States and elsewhere. If the Senate sits on the treaty indefinitely, an opportunity to start pushing the nuclear genie back into the bottle will be lost — perhaps for ever. The Senate must grasp that opportunity and ratify the Comprehensive Test Ban Treaty in 1998. □



Admission on Gulf War vaccines spurs debate on medical records

[PARIS] Pressure in the United Kingdom for an overhaul of military medical research was reinforced last week by an admission from the UK Ministry of Defence (MOD) that it had failed to heed warnings at the start of the Gulf War about the possible side-effects of a vaccine combination administered to British troops.

A report presented to the House of Commons by John Reid, the armed forces minister, confirms that a pertussis (whooping cough) vaccine was given to troops as an adjuvant for an anthrax vaccine, so that the latter took effect after seven instead of 32 weeks.

Use of the pertussis vaccine in this way was highly experimental, relying on preliminary results from MOD-sponsored research at the Centre for Applied Microbiology Research (CAMR) at Porton Down, but was done to get troops out to the Gulf quickly.

But the report reveals that the National Institute for Biological Standards and Control (NIBSC) had warned the MOD, in a fax dated 21 December 1990, that in animal studies this combination caused "severe loss of condition and weight". Although the MOD recorded receipt of the fax, no-one there could recall "or admitted to recalling the fax", said Reid.

The revelations have also lent weight to suspicion that vaccinations might account for some of the purported symptoms of Gulf War syndrome. A recent paper, by Graham Rook and A. Zumla from University College London, suggests that multiple vaccinations may cause a large change in the immune response that could result in similar symptoms to those observed in Gulf War veterans (see *The Lancet* 349, 1831; 1997).

Circumstantial evidence pointing to vaccines as a cause for the syndrome also comes from the fact that no symptoms have been reported in French forces, who were not given



Ready for action: but did haste in getting troops to the Gulf mean that short-cuts were taken?

the vaccines used by British and US troops.

If this difference is confirmed, it would be a "crucial epidemiological point", says Simon Wessely, a researcher at King's College in London who is leading a study on Gulf War syndrome. But he adds: "I'm only prepared to accept this when there is evidence that the French have looked for [symptoms]; they haven't systematically looked yet."

Whether or not the vaccine combination did in fact harm soldiers, the British admission, coming seven years after the incident, has raised new questions about the government's handling of the Gulf War syndrome controversy. And further queries arise from last week's revelations that a former defence minister gave incorrect information to Parliament on the extent of the use of organophosphate pesticides during the war.

Most observers conclude that the MOD could have done more to test the prophylaxis given to troops. They point in particular, to the fact that the NIBSC's action in checking the safety of using pertussis as an adjuvant

owed more to chance than to MOD research planning. The MOD asked the NIBSC only to verify the safety of the pertussis vaccine — which was licensed in France but not in the United Kingdom — without revealing that the vaccine was to be used as an adjuvant. But the NIBSC "deduced" that this was the purpose and "therefore decided that a check for interactions might be helpful", according to the report.

"MOD suffers from an excessive culture of secrecy; they don't deal well with outside researchers," says one academic UK researcher who has worked with the MOD on Gulf War syndrome. He describes the ministry as "hideously complicated", adding that while he has had "excellent help" from parts of it, others have given him none.

The government comes in for similar criticism over its announcement last week that it intends to fund a £2.25-million research programme into the effects of the combination of vaccines and anti-nerve gas treatments that were given to troops; many describe this action as belated. But other observers warn against excessive hindsight, and sympathize with the dilemma created by the need to protect troops when licensed products may not be available.

For example, Jack Woodall, director of the arbovirus laboratory at the New York State Department of Health in Albany, and an expert on biological weapons, describes the MOD's admitted failure to keep proper medical records as having resulted in a "missed opportunity to do a proper scientific study to see who responded to vaccines and combinations".

But Colonel Wilbur Milhous, director of experimental therapeutics at the Walter Reed Army Institute of Research in Washing-

Ariane-5 saves Europe's place in space

[PARIS] Europe's Ariane-5 launcher, which exploded on its maiden flight last year, is back on track after last week's successful launch from Kourou in French Guiana. The launcher, which cost US\$7 billion to develop, is Europe's main space project and its ticket to maintaining its autonomy in space and continuing its dominance of the commercial launch market.

A repeat of the first failure, which cost the European Space Agency \$350 million, would have plunged Europe's cash-strapped space programme into crisis (see page 8).

Ariane 5 was originally designed to carry Hermès, a planned European space shuttle that has since been abandoned. Somewhat poignantly, France last week decided to withdraw from the last relic of its Hermès activities, a joint US-European programme to build a crew rescue vehicle for the international space station. The decision reflects the aversion of Claude Allègre, the new space minister, to manned space flight, as well as the end of the country's ambitions in this area.

D.B.

ton DC, argues that carrying out such studies under deployment conditions is often unrealistic. "In Bosnia, conditions were horrendous. One guy was given a jeep and a technician and told to go out and immunize thousands of people.... just getting the vaccines delivered was a major accomplishment, much less keeping records of who got what."

Similarly, one former CAMR official claims that a ceiling imposed for political purposes on the number of UK servicemen that could be sent to the Gulf led to economies in medical research staff. "It would have been nice to have people collecting data [on the effects of prophylaxis], but it was more important to have tank forces than clever scientists. I remember there were tight constraints."

The controversy about Gulf War syndrome has sparked renewed attention in the United States to protocols associated with experimental prophylaxis and therapeutics under battlefield conditions; review panels are being set up by the Food and Drug Administration and the Department of Defense to consider the issue.

But, according to Wessely, even if practical constraints to substantially improving the conduct of research in deployment settings are accepted, much could still be done to improve post-conflict surveillance. "There is no excuse for the sluggish response of those in authority to sick servicemen coming back from war," he says.

Although military service will always create its own health problems, the United Kingdom has failed to establish "a routine system for data collection, epidemiological surveillance, and follow-up", he says. If such a system had been in place at the time of the Gulf War, he adds, it would have been much easier to look for patterns and conclude whether there was a problem or not.

One recommendation his group's study is likely to make is that the health of veterans should automatically be monitored so that unusual outbreaks of cancer, reproductive defects or deaths can be detected. At present, it is impossible to tell whether servicemen suffer from higher levels of cancer, for example, says Wessely.

Veterans need to be treated as a special group for research purposes, with particular health needs, he says. "Regardless of whether Gulf War syndrome exists or not, many veterans are bitter and disillusioned at the feeling there has been official disregard for their health."

On a wider level, the short-cut taken by the MOD in using pertussis vaccine as an experimental adjuvant for an anthrax vaccine raises the question of why the ministry had not carried out comprehensive research into developing a fast-acting anthrax vaccine sooner. A former official from Porton Down says that the West was unprepared for the threat of an enemy striking rapidly with biological and chemical weapons. **Declan Butler**

Critics claim US inquiry was 'irreparably flawed'

[WASHINGTON] The US government came under renewed heavy criticism last week for its investigations and handling of the elusive collection of symptoms known as 'Gulf War illnesses'. Two reports, one adopted by a congressional committee and one still in draft form, have highlighted research into chemicals exposure during the war.

The House of Representatives Committee on Government Reform and Oversight adopted unanimously a report that assails the Departments of Defense (DOD) and Veterans Affairs (VA) for the weakness of their efforts to establish the cause or causes of Gulf War illnesses. This failing of research, the report says, as well as lapses by the Central Intelligence Agency (CIA) and the Food and Drug Administration (FDA), amount to a government effort that the congressional committee called "irreparably flawed".

In particular, says the 140-page report, the DOD and VA resisted and then mishandled investigations into chemicals exposures as a potential source of the various symptoms experienced by some 100,000 Gulf War veterans. The departments' behaviour, it says, has been "plagued by arrogant incuriosity and a pervasive myopia that sees a lack of evidence as proof".

Among the report's recommendations are that the DOD and VA should lose their authority over further research on Gulf War illnesses to another agency able to develop a research agenda "more objectively".

Equally harsh criticism is contained in another report – still officially private – handed on 31 October to President Bill Clinton, who is expected to release and react to it within two weeks. That report, by the Presidential Advisory Committee on Gulf War Veterans' Illnesses,

praises the DOD for improving its research, but says the Pentagon "failed to pursue, acknowledge or even account for" chemicals exposures said to have been detected by US Marines who crossed an Iraqi minefield when entering Kuwait in February 1991.

A draft of this presidential report was obtained by the *New York Times*, which printed excerpts last week. These called the DOD's failure to investigate the minefield incident fully "highly damaging to DOD's credibility". The news account quotes the report as stating that "the committee perceives that public mistrust about the Government's handling of gulf war veterans' illnesses has not only endured, it has expanded".

The VA and DOD declined to respond officially to the presidential report until it is made public. But Captain Tom Gilroy, a DOD spokesman, calls the February 1991 minefield incident "old news", and points out that since July, the DOD has had posted on its website a full account of the incident, as well as a related chapter from a top-secret report compiled for DOD by the Mitre Corporation of Bedford, Massachusetts. (The website address is <http://www.gulfink.osd.mil>.)

Joyce Lashof, chair of the Presidential Advisory Committee and a professor emerita of public health at the University of California, Berkeley, also declined to comment on the substance of her committee's report. But she said any exposure suffered by Marines in the February 1991 incident was "very low level, subclinical", and that the committee stands by the position it took in a January 1997 report: "that available scientific evidence does not indicate that long-term health effects occur in humans following low-level exposure to chemical

warfare agents".

The congressional report argues otherwise, calling the circumstantial evidence for such a link "overwhelming". But it says proof may be lost forever because of the DOD's and VA's handling of the research. The departments, says the report, prematurely ruled out toxic exposure as a cause of Gulf veterans' symptoms, assuming that unless there was an immediate acute reaction, exposure could not have caused long-term symptoms. As a result, "federal research strategy has been blind to promising hypotheses", the report states.

Citing a 1997 report by the General Accounting Office, it notes, for instance, that the government funded multiple studies of the role of stress in Gulf War illness, but not until forced by legislation in 1996 did it begin funding studies on the effects of low-level chemicals exposure. Three such studies had previously been denied funds.

Responding to the congressional report, Gilroy of DOD says that "in the last year we have done a lot of work and I don't know if any of it is reflected in [the House report]". The work referred to is that done by the DOD's Office of the Special Assistant for Gulf War Illnesses, set up in November 1996 to oversee DOD research.

And the VA says the report fails to "recognize the openness and public accountability of the existing research effort". Also, says a VA spokesman, Terry Jemison, VA research undergoes vetting to protect it from political influence.

Despite all these protestations, the lead author of the congressional report, Congressman Christopher Shays (Republican, Connecticut), has promised to draft legislation early in 1998 removing investigative powers from the DOD and VA. **Meredith Wadman**

Fusion results 'confirm the scientific case for ITER'

SOURCE: JET

[LONDON] Plans for a next-generation experimental fusion reactor have been boosted by the latest research with mixed deuterium-tritium (D-T) fuel on the Joint European Torus (JET), based near Oxford, England.

The experiments, carried out over the past few weeks, have as anticipated set new world records for fusion power, fusion energy, and the critical 'fusion amplification factor' or Q — the ratio of fusion power produced to net power input.

Equally significantly, the experimental results have been broadly in line with predictions — a crucial requirement given that the design of JET is close to that of the proposed International Thermonuclear Experimental Reactor (ITER), which its supporters hope will be built early in the next century.

"There have not been any breakthroughs in the sense of something unexpected being obtained, but the physics models have met expectations, giving us confidence that the predictions for ITER are sound," JET's director, Martin Keilhacker, said last week.

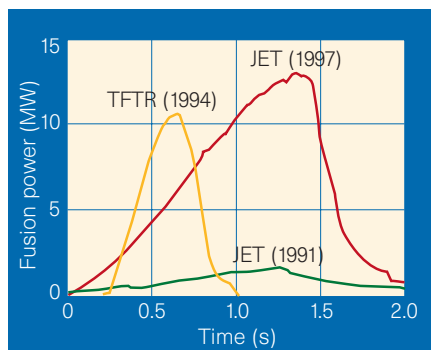
But Keilhacker and others are increasingly nervous about the funding prospects for ITER, particularly in the light of reports that the United States, one of the four planned supporters of the project — the others being the European Commission, Japan and Russia — is becoming lukewarm. "The political climate at present is not very favourable to megabillion projects," says Keilhacker.

The previous record for fusion power, 10.7 MW, was set by the Tokamak Fusion Test Reactor (TFTR) at Princeton, New Jersey, in 1994. The latest experiments with JET, which is financed largely through the European Commission, achieved a peak power of 13 MW with the power remaining at more than 10 MW for half a second (see figure).

One particularly reassuring result has been the demonstration that, as predicted, a significant contribution is made to the plasma power balance by the α particles produced in the D-T reaction, in that the temperature and energy content of the plasma were consistently higher with higher levels of α -particle heating.

"We had some doubts until we had these results whether 300 MW [the planned heating power to the plasma of ITER] would be sufficient to contain the plasma. Our results with the D-T experiments have confirmed our predictions that the α particles are trapped and heat the plasma," says Keilhacker. "Thus the process by which ignition and thermonuclear burn will occur in ITER has been experimentally confirmed."

Another source of satisfaction for the JET



Pushing the envelope: JET's output has overtaken the previous record set by Princeton's TFTR.

engineers has been the fact that, compared with previous experiments using pure deuterium, the heating power needed to access the so-called high confinement or 'H' mode, which considerably improves the thermal insulation, was significantly lower using both D-T mixture and pure tritium.

"This is a very significant result, since it predicts a 50 per cent reduction in the power needed to access conditions with high thermal insulation in a pure tritium plasma, and a 25 per cent reduction in the power needed to maintain such conditions during high fusion power operation of ITER," says Keilhacker.

But if the physics prospects for ITER look good, optimism about the outcome of political negotiations is in shorter supply. The team drawing up the detailed design for ITER, which reports earlier next year, is expected to confirm the \$6-billion cost that was planned. But a firm decision on whether to proceed — and if so on what site — has now been put off for three years (see *Nature* 388, 119; 1997).

JET officials say there appear to be no technical obstacles to construction, and — in apparent contrast to those responsible for the US fusion programme in Washington — are keen to move quickly to the construction phase with a minimum loss of momentum.

In this context, they are reacting cautiously to proposals from some senior US fusion physicists that the United States should divert some of the money that might otherwise be dedicated to ITER to support experimental work at JET (see *Nature* 369, 769; 1997).

"There is certainly scope for more US involvement [in JET] and we already have very good collaboration with US laboratories," says Keilhacker. "But there is the political problem that we in Europe are convinced that the fusion community is ready to build ITER, while the United States — partly for budgetary reasons — is not." **David Dickson**

Monsanto links up with Millenium on genome research

[BOSTON] The agricultural and pharmaceutical company Monsanto last week signed a research agreement with the genomics company Millennium Pharmaceuticals that is being hailed as one of the largest deals in the field of genomics.

Millennium will transfer its proprietary technologies in genomics, gene sequencing and bioinformatics to a new subsidiary formed by Monsanto for developing plant and agricultural products.

The subsidiary, which is expected to recruit at least 100 scientists during 1998, will be located near Millennium's headquarters in Cambridge, Massachusetts. Monsanto, which produces agricultural products, food ingredients and pharmaceuticals, is based in St Louis, Missouri.

The agreement will also provide Monsanto with non-exclusive rights to Millennium's genetics process technology for research in other areas of life sciences, including pharmaceuticals and nutrition.

Millennium, in turn, will receive \$118 million in advance and an additional \$100 million over the next five years if mutually agreed research objectives are realized. The company's vice-president of business development, Alan Crane, calls the deal "the biggest partnership we've entered into and our first relationship outside human health care". Millennium has already entered into collaborations with nine other corporate partners including Bristol-Myers Squibb, Eli Lilly, Hoffmann-La Roche and Pfizer.

Monsanto is interested in pursuing several agricultural applications through the deal, including the introduction of new herbicides and pesticides, the development of genetically engineered plants and seeds, and the application of genomics tools to the process of "directed breeding".

Monsanto's president, Hendrik Verfaillie, says Millennium's "broad, integrated technology platform" will "greatly increase the speed and precision with which we can analyze new product leads".

The partnership fits in well with Millennium's strategic mission. According to Crane, the companies "share a similar vision of the potential of technologies like genomics to revolutionize the life sciences". **Steve Nadis** ● Sequana Therapeutics, the gene discovery company based in San Diego, California, announced on Monday that its stock has been acquired by Arris Pharmaceutical Corporation in a deal said to be worth approximately \$166 million. The new company, which will bridge the range from genomics to clinical application, will operate under the name of AxyS Pharmaceuticals.

NIH heads for 7.1 per cent budget growth

[WASHINGTON] A joint spending committee of the US Senate and House of Representatives last week approved a 7.1 per cent funding increase for the National Institutes of Health (NIH) — to \$13.65 billion — in the fiscal year 1998, in a bill reconciling differences between the two legislative bodies.

Such a bill would normally move speedily to the floor of the House and then the Senate for final passage. But as soon as the politicians had signed off the massive bill last Thursday (30 October) it became captive to a sharp feud between Republican factions in Congress, and between Congress and the White House, over an unrelated education measure it contains concerning national examinations in reading and mathematics.

As a result, Newt Gingrich (Republican, Georgia), the Speaker of the House, barred the bill from proceeding to the House floor. At the beginning of this week, it was unclear whether Gingrich would order the reopening of negotiations on the \$269-billion bill, which funds the Departments of Labor, Health and Human Services, and Education.

According to John Porter (Republican, Illinois), who chairs the House appropriations subcommittee responsible for NIH funding, a likely alternative is that the bill will

be bundled together with several other outstanding spending bills containing equally controversial measures. These would go to a single vote on the House floor, perhaps as early as this week.

As it contains so many spending provisions, such an 'omnibus' bill would be intended to force politicians to "swallow hard and pass it", says David Moore, a lobbyist for the Association of American Medical Colleges. It would also make it more difficult for the White House to veto the larger bill.

Whatever the outcome of the impasse on examinations, Porter said last Friday that he was confident that the NIH increase is secure. "Unless there is a complete breakdown, which I don't contemplate, NIH will receive a 7.1 per cent increase for the next fiscal year," he said.

The earlier House version of the appropriations bill had contained a 6 per cent increase for NIH, while its Senate counterpart had sought 7.5 per cent. The committee's bill also provides the NIH with authority to spend \$100 million on research into Parkinson's disease in 1998, and allocates nearly \$17 million to begin construction of an AIDS vaccine research centre — reduced from an earlier House figure of \$26.1 million.

In another change, Tom Harkin (Democrat, Iowa) succeeded in boosting by two-thirds the budget of NIH's controversial Office of Alternative Medicine (see *Nature* 389, 652; 1997). The bill as approved by the committee provides \$20 million to the office. Its current budget is \$12 million and NIH's director, Harold Varmus, had called for that to be cut to \$7.5 million.

Also last Friday, Porter repeated his intention to seek an additional \$2.5–\$3 billion for NIH in 1999 as part of a drive to double its budget over the next five years. The increase could be funded, he said, by government revenues that had not been anticipated during the drafting of a balanced budget agreement that set tight caps on such new spending.

"There may be some very strong rationale for varying to some degree the balanced budget agreement to allow us to have this initiative on research," said Porter. "Given the good economy, it's a good time to propose it."

But he added that he was willing to drop the idea if NIH leaders feel that such a large, quick increase would overwhelm their capacity to deal efficiently and effectively with it. "We don't want to give them money they can't use wisely and for good research," he said.

Meredith Wadman

Canada 'still has a long way to go' in effective control of acid rain

[MONTREAL] Large areas of North America will have to cut sulphur dioxide emissions by an additional three-quarters if they are to control acid rain — a problem many believed to be a thing of the past.

A report published last month by the Acidifying Emissions Task Group, set up by the Canadian government, concludes that this extra cut in emissions is needed in the Canadian provinces of Ontario and Quebec and in midwest and eastern parts of the United States.

The task group, made up of representatives of government, industry and environmental and health groups, was responding to a 1994 request from Canadian environment and energy ministers for a strategy to mitigate the environmental and health effects of acid rain.

"The assumption has been that the problem is solved," says Wayne Draper, a senior official in the federal environment department who acts as joint chairman of the group. "In reality, we took only a first step and we still have a long way to go."

That first step, however, has already halved sulphur dioxide (SO₂) emissions in seven eastern provinces. Begun in 1985, the Eastern Canada Acid Rain Program's goal was to reduce wet sulphate deposition to no more than 20 kg per hectare per year. A



smaller Sulphur Oxide Management Area in the southeast was to cap SO₂ emissions at 1.75 million tonnes a year starting in 2000. Last year, emissions were already 29 per cent below that cap.

The United States has cut emissions by 30 per cent since 1980, and by 2010, when its acid rain programme is fully implemented, emissions should be down by 40 per cent. But this will still leave US emissions about five times greater than those from Canada and responsible for more than half the total acid deposition in eastern Canada.

"Even in 2010, with full implementation

of the Canadian and US programs, almost 800,000 square kilometres in southeastern Canada — an area the size of France and the United Kingdom combined — will receive harmful levels of acid rain; that is, levels above critical load limits for aquatic systems," says the report, *Towards a National Acid Rain Strategy*.

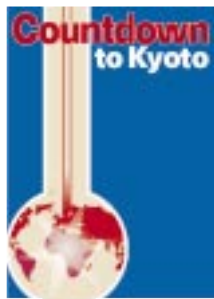
Tests show that, while sulphate levels are declining in most lakes in Ontario and Quebec and are stable in the Atlantic region, acidity levels remain high. Only 33 per cent of 202 lakes tested in Ontario, Quebec, Nova Scotia and Newfoundland between the early 1980s and the mid-1990s contained less acid, 11 per cent had more and 56 per cent showed the same levels.

"Acidified lakes pose a serious threat to biodiversity in eastern Canada," says the report. "Loons do not breed as easily or lay as many eggs, and there are fewer species of clams and crayfish — an integral part of the food chain."

The group recommends negotiating with the United States for further reductions, and reviewing related science, research and monitoring programmes. The report will go to a meeting of energy and environment ministers later this month, who will try to propose a national acid rain strategy.

David Spurgeon

World Bank fund finds allies and sceptics



[LONDON] The World Bank has offered to set up and manage a fund to help combat global warming. Under the scheme, known as the Global Carbon Initiative, developed countries would pay for low-cost energy-efficient

projects in developing countries. The ensuing greenhouse gas savings could then be credited to the donors' legally binding target of greenhouse gas emissions under a system known as 'joint implementation'.

The bank aims to collect US\$100 million, and has begun to gather support for the idea. Switzerland, Norway, Finland, Sweden, the Netherlands and several private corporations have already expressed an interest in getting involved. The fund would invest in carbon-abatement projects in the bank's client countries. The bank says it has access to US\$6-billion-worth of projects, encompassing energy, industry, transport and land use.

The bank would act as investment manager and decide on the portfolio of projects. Apart from obtaining carbon or other greenhouse credits, the advantage for investors, it says, is that projects would be considerably cheaper than comparable ventures in developed countries. Developing countries would benefit from the creation of jobs, technology transfer and the spread of energy-efficient technology.

But developing countries and some non-governmental groups have yet to be convinced of the scheme's merits. Critics fall into two camps: those who disagree with both the concept and the World Bank's involvement, arguing that the idea represents the interests of Western investors and governments; and those who agree with the idea of joint implementation, but are uncomfortable about the bank's potential role as monopoly provider.

Anil Agarwal, director of the Centre for Science and Environment in New Delhi, is in the first category. He was asked by the Indian government to comment on the proposals, and plans to recommend their rejection, partly because the scheme assumes two standards of living: a lower one for developing countries and a higher standard for the developed world.

In an editorial in the latest issue of the centre's fortnightly magazine *Down To Earth*, Agarwal writes that the bank wants to "buy the rights of present and future generations of Indians to the common atmosphere. And would sell their rights so cheap that even the American Indians who sold New York for a few beads, or the Russians who

sold Alaska for a few dollars, would be rich in comparison."

Some others who are broadly in favour of joint implementation are also nervous. Kilaparti Ramakrishna, director of the science and public affairs programme at the Woods Hole Research Center in Massachusetts, says that for many countries the concept is not at issue. "The problem is with the World Bank, which has a major credibility problem in many countries."

Ramakrishna says some members of the Group of 77 developing countries are unhappy about the bank being given the authority to decide in which countries to invest. He says they would prefer the flexibility of dealing directly with donor countries and private corporations.

Ramakrishna adds that developing countries are also concerned that the new fund could interfere with the planned replenishment by governments of the Global Environment Facility (GEF), a United Nations environmental fund that is administered through the World Bank.

But Johannes Heister, a World Bank economist in Washington DC who helped to design the initiative, says such concerns are misplaced. He points out that the fund would not begin until next year, well after

the GEF's replenishment.

He says that donor countries have promised that their investments would not affect how much they will give to the GEF. The amount pledged, he adds, is minimal compared to the expected \$2-\$3 billion GEF replenishment.

Heister argues that the bank will simply be providing a service that will be needed if joint implementation is agreed on in December at the annual conference of the international climate convention in Kyoto, Japan. He says the bank is uniquely placed to obtain sufficient finance, which is essential to the success of joint implementation. "It is an open market. Anyone can get involved."

He adds that the bank has long experience of dealing with governments and the clout to obtain private-sector finance. Transaction costs for World Bank projects are also significantly lower than those incurred through bilateral deals.

But the bank recognizes that its plan would work only if the Kyoto conference results in agreement on legally binding greenhouse gas emissions targets — and if such a treaty includes the principle that countries can receive greenhouse credits by paying for emissions reductions in other parts of the world.

Ehsan Masood

Kyoto 'dress rehearsal' ends in deadlock

[LONDON] Two weeks of negotiations towards a treaty on greenhouse gas emissions ended in Bonn last week as they had begun — in deadlock. One observer described the meeting as "little more than a dress rehearsal" for the climate convention's annual conference next month in Kyoto, Japan.

The Bonn meeting's chairman, Raúl Estrada-Oyuela of Argentina, remarked that negotiations seemed locked in a 'time warp' with delegations repeating positions he had heard three decades ago. Demands from the European Union and developing countries for a 15 per cent reduction in emissions below 1990 levels by 2010 continue to be opposed by the United States, which favours stabilizing emissions at 1990 levels.

The United States

continues to insist that commitments by developing countries must be part of the Kyoto protocol. Australia refuses to sign any legally binding protocol, despite earlier signs that it might do so (see *Nature* 389, 893; 1997). Japan's compromise proposal of a 5 per cent reduction between 2008 and 2012 now seems the only realistic option on the table.

Estrada successfully resisted US attempts to include developing country commitments in the formal text of the protocol. These will now be discussed separately at Kyoto. But the United States angered many delegates by suggesting on the last day of the meeting that armed forces should be exempt from an emissions reduction protocol.

The draft document that now goes to Kyoto contains 10 articles, two annexes and an attachment, most of



Spot the dinosaur: Greenpeace protests at Bonn meeting.

which are full of blocks of square brackets — denoting text that has yet to be agreed. The most heavily bracketed part relates to the timing and size of greenhouse gas reductions, whether reductions should be calculated according to a 'basket' of gases or individual gases, and whether reductions should be calculated annually or as an average over a period of years.

E.M.

Space science projects may be grounded

[MUNICH] Leading European space scientists warned last week that Horizons 2000, the long-term science programme of the European Space Agency (ESA), could disintegrate unless a cap on its budget imposed by European governments two years ago is lifted in 1999.

The warning has come from the eight independent scientists who make up ESA's Space Science Advisory Committee (SSAC), and is based on a 'worst-case' scenario of what might happen if member states refuse to increase funding.

The decision to freeze the agency's science budget for three years was taken at ESA's last ministerial meeting in Toulouse in 1995. At that meeting, Europe's space ministers also agreed to consider extending the freeze to 2000 at their next scheduled meeting, which will probably be held in Bruges, Belgium, next June (see *Nature* 377, 667; 1995).

The loss of purchasing power resulting from the three-year freeze — estimated at 9 per cent — forced ESA to modify its Horizons 2000 programmes, which had been designed to run to 2009, assuming an index-linked budget to 2002 and an annual 5 per

cent increase thereafter. The revised plan, whose concept was approved by ESA's Science Programme Committee in September, maintains the regular ECU650-million (US\$740-million) 'cornerstone' missions. But it abandons the ECU300-million medium-sized missions in favour of smaller, 'flexible' missions which can be planned at short notice and cost only ECU175 million, and low-cost technology-driven SMART missions (see *Nature* 389, 218; 1997).

According to this plan, four cornerstones, three SMART missions and seven flexible missions could be launched in the next 12 years. But the SSAC says it will not be possible to retain all these elements without resumption of a fully index-linked budget after 1999 and annual 5 per cent increases after 2002.

Last week, the committee put forward two new calculations based on worst-case scenarios. In the first — restoration of index-linking, but failure to increase the budget after this — the cornerstone missions could be launched but the SMART missions would have to be abandoned.

In this scenario, the idea of flexible

missions — based on the 'better, faster, cheaper' missions of the US National Aeronautics and Space Administration (NASA) — would, according to the SSAC, have to be abandoned, as their infrequency would produce insufficient data to keep the space science community alive. NASA scientists have also recently expressed doubts about the scientific value of 'lean missions' (see *Nature* 389, 899; 1997). So, instead of flexible missions, three medium-sized missions would be launched.

In the second scenario, which the SSAC refers to as the 'disaster scenario', the budget would remain capped at its 1999 level for the duration of the Horizons plan. If this were to happen, only three cornerstone missions and two medium-sized missions could be launched.

Both scenarios would mean delays to either the planned mission to Mars, Mars Express, or the infrared cornerstone FIRST mission, which is to be combined with the microwave background mission Planck.

Lodewijk Woltjer, outgoing chairman of the SSAC, says that, if the budget continues to decrease in real terms, "the Horizons 2000 programme will have to be effectively abandoned — Europe will just have to do the occasional mission".

ESA's Science Programme Committee will formally respond to the SSAC document this month. Informal responses indicate concern, and some delegations hope that their ministers will be persuaded to lift the freeze. Italy, the third largest contributor to ESA's budget, has made provision for this in its five-year space plan (see *Nature* 389, 651; 1997). But some countries, including Britain, Germany and Sweden, are likely to insist on a continuation of the budget freeze at the ministerial meeting next year.

Germany, the largest contributor to ESA, would like to support an increase, says Norbert Kiehne, director of the space programme of the German air and space agency DLR. But its hands are tied because there will be no increase in its overall space budget in the next few years, and priority is being given to commercial applications of space activities, according to a new government policy document.

Some believe there are other alternatives to the disaster scenario. Kiehne says that ESA should continue to improve its efficiency, to do more missions for less money. Paul Murdin of the Particle Physics and Astronomy Research Council, the UK delegate, agrees. The original estimates for the replacement for Cluster — the solar mission that was destroyed when the first Ariane 5 launch exploded last year — were more than halved when the Cluster team and ESA were put under pressure, he says.

Alison Abbott

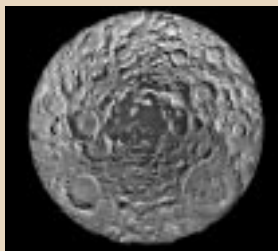
Millennium Moon mission seeks private funds

[MUNICH] Plans being drawn up by the European Space Agency (ESA) to send a robotic expedition to the south pole of the Moon to celebrate the millennium have received a cool response from the agency's Space Science Advisory Committee (SSAC).

The Euromoon project was chosen from two 'millennium' ideas presented to ESA's council last December. The other idea — the Star of Tolerance, a pair of orbiting balloons that would reflect sunlight back to Earth — was vetoed because of complaints that it would pollute the sky.

Euromoon was dreamed up by Wubbo Ockels, a former astronaut who now works at ESA's technology development centre, ESTEC, in the Netherlands. The mission would place a lander on the rim of the large Aitken crater at the lunar south pole.

The perpetual sunlight at that point — known as the Peak of Eternal Light — could



Bottom up: the lunar south pole, seen in a composite image.

provide constant solar power to the lander. Robots would gather samples from the crater, which is thought by some to contain ice.

The mission is ambitious, as a space vehicle would have to land with great precision on an area no larger than 100 metres square. Costs would be kept down to around ECU300 million (US\$342 million) by using off-the-shelf technologies and a small team of young scientists. The plan assumes support from ESA's own technological and operational facilities.

Euromoon, which is already registered as a

trademark, must seek private sponsorship as ESA is unlikely to use its own funds. The SSAC was asked to help develop a science programme for the mission. But at a meeting last week the panel expressed concern about the burden it would place on its overstretched budget, and the fact that the mission has not passed through the standard peer-review processes.

The committee has asked its solar systems science working group for further advice. If the mission is to be launched in time for the astronomical millennium — which starts in 2001, rather than 2000 — Euromoon must secure financing quickly.

European space industries have indicated willingness to provide resources for the project. But a major bid has yet to be launched to persuade other industries to inject cash in return for vaguely defined commercial gains and advertising prospects.

A.A.

Dutch reforms raise alarm in universities

[MUNICH] The Dutch government plans to increase central control over research spending by transferring Df1500 million (US\$250 million) of university funds to the national research council, the NWO. The rules of the council — at present an independent agency — will be changed to allow the government to decide how the new money is to be spent.

The transfer would double the current NWO budget, and reduce that of the universities by 20 per cent. At the same time, the government wants to separate the administration of the NWO's 14 research institutes from its grant-giving activities, to avoid potential conflict of interest.

The plans, announced in September by Jo Ritzen, Minister of Education, Culture and Science, have been generally welcomed by the NWO. But they have been greeted with alarm by the universities, unwilling to lose control of the distribution of research funds.

Some of the NWO's discipline-orientated advisory groups, known as 'foundations', which recommend how grant money should be distributed, have also argued strongly against creating a new organizational structure for NWO research institutes, many of which are closely integrated into universities.

"It would isolate institutes from universities," says Huup Eggens, head of the natural



Ritzen: mixed reaction to research proposals.

sciences foundation (FOM), which represents the high-energy physics community. Others express concern about the vagueness of Ritzen's plan to split the functions of the NWO. But they are optimistic that it could bring improvements. "The transition period has not been thought through in detail," says Harvey Butcher, head of the Netherlands Foundation for Research in Astronomy and director of the NWO's institute for astronomy in Dwingeloo. "But there is room for improvement in the way our institutes are run, provided changes are carried out in a professional way."

Dutch natural scientists also complain that the announcement of Ritzen's plans has accelerated the NWO's continuing internal restructuring process, which is intended to simplify its decision-making. The number of foundations has been cut from 34 to 20 over the past four years, but the NWO plans to eliminate this whole level of decision-making, and transfer its activities to the area councils which are responsible for much broader

disciplines. Eggens says, however, that taking away control from the experts "will have a deleterious effect on the quality of research".

Predictably, universities are upset at the prospect of losing control over a substantial proportion of their budgets. A spokesman for the Association of Universities, Jan-Willem Vos, says that such a drastic change could force redundancies among university staff. Vos says the universities feel cheated because they agreed only six months ago that the NWO would distribute Df1200 million from the universities' budget to the ten university research groups judged, through open competition, to be working in research areas designated as strategic priorities by the NWO.

But Ritzen says that the transfer would be carried out over a few years, to avoid disruption within universities. He says his plan would increase competition between universities and allow the government to have a role in setting research priorities based on industrial and social needs that is more in line with its European Union neighbours.

Hein Meijers, a spokesman for the NWO, says the organization is happy for the government to help define priority research areas — provided that the proportion of its funds used for strategic research does not exceed 50 per cent.

Alison Abbott

South Africa increases central control of higher education

[CAPE TOWN] South Africa's parliament last week approved a major reform of the country's higher education system, despite concern about language policy and centralized control.

The bill's most important consequence will be the creation of a Council on Higher Education to accredit degree programmes, allocate funds to universities and technikons (advanced technical colleges) on a three-year cycle, and audit quality assurance. The council will also advise on the intractable problem of student financial aid.

The Council on Higher Education will comprise 14 voting and 6 non-voting members, but its real power is vested in a five-person executive committee that can act on the council's behalf. Although the full council can revoke the executive committee's decisions, action taken as a result of an executive committee decision before its revocation will remain valid.

The government hopes to appoint the members of the council by the end of the year. The first six months of next year are likely to be spent allocating places for programmes to institutions, and deciding the financial value of subsidies for different degree programmes.

This process is likely to be controversial if

the council pursues the government's stated aim of redressing the present huge bias in favour of subsidized places in arts and social science — as opposed to natural-science-based disciplines — as this bias is most prevalent in the former black universities. But the process should be helped by the government's decision to provide additional earmarked funds to correct past imbalances.

Earmarked funds for research are provided for in the subsidy formula of the higher education white paper's final version — unlike an earlier version (see *Nature* 387, 327; 1997) — which was officially published in April. This provision allows the minister to allocate funds for preserving and strengthening areas of research excellence, as well as for developing new ones.

Specifically, the white paper emphasizes that enrolment levels in doctoral programmes are not only too low but are also biased in terms of race and gender. The new system is intended to support postgraduates in fields where institutions have demonstrable research training capacity. Provision is also made for institutions to apply for earmarked funds to enhance the infrastructure necessary to support expanded postgraduate training.

The right-wing Freedom Front was one of

the parties opposing the bill — the party is concerned about the future of Afrikaans as a medium of instruction in universities. Freedom Front students in the public gallery disrupted the bill's debate in parliament.

They were protesting against a provision allowing the education minister to decide the language policy of tertiary institutions on the advice of the Council on Higher Education.

The National Party and the Democratic Party also voted against the bill, the latter on the grounds that it gives the government increased control over higher education institutions.

Walter Claassen, vice-rector for academic affairs at the Afrikaans-medium University of Stellenbosch, says this centralization of power is "very worrying". While acknowledging the government's desire to change the system in accordance with national goals, he feels that "this should not be done in such a way that the autonomy of universities is threatened".

But Claassen describes the decision to allow the new higher education council to allocate funds to universities and technikons on a three-year cycle as a "very promising development, as it could provide a more secure basis for institutions in planning enrolments".

Michael Cherry

Mars Surveyor faces 12-month delay in reaching final orbit

[WASHINGTON] The Mars Global Surveyor will resume lowering its orbit around the planet this week, and project managers expect that most of the mission's scientific goals can still be met. 'Aerobraking' of the Surveyor — using drag from the thin Martian atmosphere to drop it to a lower, more circular orbit — was suspended on 12 October because of an unexpected wobbling of the spacecraft's solar panels (see *Nature* 389, 772; 1997).

Engineers now think they understand the problem well enough to resume aerobraking, but will proceed at a slower, more cautious pace. Systematic mapping of the planet had been expected to begin next March; now the spacecraft will take eight to 12 months longer to reach the proper mapping orbit.

Scientists and engineers will spend the next few weeks deciding which orbits could best satisfy competing demands by different onboard instruments for optimal viewing angles, distance from the surface, and other factors.

The mission manager, Glenn Cunningham of the Jet Propulsion Laboratory in Pasadena, California, said last week that the spacecraft should still be able

to meet its goal of mapping both the southern and northern hemispheres of Mars.

East and west Germany equal on institutes

[MUNICH] The heads of Germany's 16 *Länder*, or states, have approved the founding of a Max Planck Institute for Evolutionary Anthropology in Leipzig. This increases to 18 the number of research institutes that the Max Planck Society has set up in eastern Germany since the country's reunification. The distribution of Max Planck institutes in east and west is now about equal.

Launch failure impact on Amazon project

[SÃO PAULO] The launch of the first prototype of Brazil's VLS rocket ended in failure last Sunday, destroying the SCD-2A data collecting satellite that it carried. The loss of the satellite will affect the collection of environmental and climate data from about 200 automatic stations throughout the country.

It had been planned to more than double the number of such stations, especially in remote parts of Amazonia. But the loss of the satellite means that the only available satellite to gather the data beamed up is the first of the series, SCD-1. This was launched in 1993

by an American Pegasus rocket and is already beyond its planned life expectancy.

The four-stage VLS rocket, developed by the Brazilian Air Force's Institute of Space and Aeronautics, is designed to launch satellites of up to 300 kg in low altitude orbits. But the prototype, launched from the Alcantara space centre in northern Brazil, had to be destroyed after 65 seconds of flight because one of the first stage's four motors failed to ignite.

German physicians 'want more data on genetics'

[MUNICH] Fewer than a fifth of Germany's general practitioners are aware of research in medical genetics that is likely to be relevant to medical practice in the near future. But they are keen to know more, and most are convinced that genomics research will have direct benefits for medicine, according to a survey by the Max Planck Institute of Molecular Genetics in Berlin and the London School of Economics.

Despite this enthusiasm, and the fact that doctors frequently discuss genetics problems with patients, there is no push to educate German general practitioners about genetic counselling, says Christopher Weber, who initiated the survey. With new genetic tests appearing on the market, he says there is an urgent need for an educational

programme and advisory support from trained geneticists.

Warning on danger of more forest fires

[LONDON] Tighter control of the use of 'slash and burn' techniques for land clearance are needed to avoid a repeat of the forest fires that swept Indonesia in September and covered the rest of Southeast Asia in smoke, according to a report issued jointly by the Nairobi-based International Centre for Research in Agroforestry, and the Center for International Forestry Research based in Bogor, Indonesia.

The report says that an outright ban on slash and burn methods will be difficult to enforce. Instead, it suggests a ban on fires in peat swamps, which can smoulder underground for months. The report adds that caution should be exercised before allocating peat lands for agricultural development.

'Irregular' satellites found around Uranus

[WASHINGTON] Two 'irregular' satellites have been found orbiting Uranus, bringing the planet's total number of 'moons' to 17. The satellites, which orbit at distances of 6 million and 8 million miles respectively from

Uranus, were first observed in September by Philip Nicholson and Joseph Burns of Cornell University, Brett Gladman of the University of Toronto and J. J. Kavelaars of McMaster University in Ontario.

After subsequent observations by telescopes in Hawaii and New Mexico, the discovery was confirmed last week by the Central Bureau for Astronomical Telegrams at the Smithsonian Astrophysical Observatory in Cambridge, Massachusetts. The two 'irregular' satellites — designated as such because of their eccentric and inclined paths — are the first to have been discovered around Uranus; Jupiter has eight, and Saturn and Neptune have one each.

WHO group warns on xenotransplantation

[PARIS] A group of scientists, lawyers and bioethicists brought together last week by the World Health Organization has warned that xenotransplantation — the transplantation of animal parts to humans — could result in the emergence of new human diseases.

The group concluded that more research is needed to establish the ability of animal infectious agents to cross the species barrier to humans. It recommended that the issue of the safety of both recipients and the wider population should be addressed at national and international levels, and called for

greater debate about ethical aspects of the technology. "Public acceptance should not be assumed," the group warned.

Society to raise profile of Indian astronomy

[NEW DELHI] India's astronomers have formed the Indian Society for History of Astronomy in order to increase international recognition of the country's contributions to astronomy over the past 5,000 years. M. S. R. Ansari, formerly professor of astronomy at the Aligarh Muslim University, has been elected first president of the society. Its headquarters will be at the Birla Science Centre in Hyderabad.

Crop scientist to head rice research body

[LONDON] The International Rice Research Institute in Los Baños in the Philippines has called a veteran agricultural research administrator out of retirement in California to take over as interim director-general, to replace the departing George Rothschild.

Robert Havener, a crop scientist who has held senior posts in several centres linked to the Consultative Group on International Agriculture Research, will take over in January. He will stay in the post until a permanent director-general can be found.

Institutes top UK science league table

Sir — Recent analyses have compared the scientific output of countries according to the number of papers published and their citations^{1,2}. These show that the United Kingdom scores highly in many areas.

Although there are reservations about the use of citations, they appear to be broadly accepted for assessment exercises and they no doubt influence policy decisions. Thus there is interest in establishing which sectors of the research base contribute to this success. For example, May¹ has speculated that the reason for the United Kingdom's relative success is because a higher proportion of its research is done in universities, although the opposite view has also been expressed³.

Previously published studies do not separate out the contribution of research council institutes to the United Kingdom's total research output. Institutes mostly concentrate on biological and biomedical research, areas in which the United Kingdom was identified^{1,2} as being strong, and much less on areas such as chemistry, in which the United Kingdom is not in the top five in impact¹.

To establish the contribution of a group of 15 institutes sponsored by, or related to, the Biotechnology and Biological Sciences Research Council (BBSRC) and listed in Table 1, the Science Policy Research Unit (SPRU) at the University of Sussex surveyed publications in all 154 fields defined by the independent Institute for Scientific Information (ISI)⁴. The publications in the SPRU database are derived from that used by ISI⁵ but include only refereed papers, notes and reviews. SPRU has verified that papers are allocated to their correct institution^{6,7}.

The results in Table 1 compare the figures from the institutes with those from the United Kingdom as a whole, the total university sector, and a group of 15 universities that did best in fields relevant to the institutes in a recent ISI analysis⁵.

The relative citation impact (RCI) for the papers in the fields in which the institutes publish most has been calculated for the different groups. In 14 out of the 22 fields, in which the institutes published more than 300 papers, the RCI for the institute papers is higher than for those

from any other group. These results remain broadly similar when the total citations from 1981 to 1994 are used in a three- or five-year window of citations.

To estimate international competitiveness, SPRU compared the citation rates for the same fields, using a three-year window, with those for the world, the United States, Canada, Europe as a whole, Germany and France. In all but three fields, the institute average citation rate of papers published between 1981 and 1992 was greater than those from any other country or group of countries.

Relative output from different organizations is difficult to assess because of their varying roles. For example, the papers in Table 1 represent less than half of the total publications of the institutes. But Krebs⁸ calculated that the non-medical university sector produced 8.26 papers per researcher, compared with 8.43 for institutes; and other sectors had lower figures.

In conclusion, the results presented here suggest that the quality of the refereed papers from the selected institutes are competitive with those from other sectors of the UK research base and with those of other countries. Other studies^{6,7} (J. S. Katz *et al.* in preparation) suggest that this conclusion can be extended to other institutes and similar organizations. So the scientific output of the United Kingdom, as measured by citations, is enhanced by papers from research institutes. The institute organization and environment does not inhibit the production of high-impact research.

This letter was prepared with the help of the directors of the institutes in Table 1. The table was produced with the assistance of Dr Sylvan Katz and his colleagues at SPRU. I am grateful to them all for their contributions.

Ben Mifflin

IACR-Rothamsted,

Harpenden,

Herts AL5 2JQ, UK

e-mail: ben.mifflin@bbsrc.ac.uk

Table 1 Relative performance of sections of the UK science base in citation impact

Science publishing sub-field ⁴	Institutes total papers	Institutes % of UK total papers	Institutes RCI	15 universities RCI	All universities RCI	Total UK C/P
Plant Sciences	3207	26.11	1.08	1.13	1.03	8.72
Agriculture	2789	42.78	1.15	0.93	0.88	5.03
Biochemistry & Molecular Biology	2185	6.54	1.17	1.03	0.88	21.32
Veterinary Science	2026	22.65	1.44	0.83	0.85	5.30
Agriculture, Dairy & Animal Science	1853	50.19	1.20	0.95	0.84	6.17
Microbiology	1248	11.64	1.24	1.05	0.92	10.76
Food Science & Technology	1227	29.77	1.48	0.74	0.72	5.97
Biology	1218	17.57	0.58	1.16	1.05	11.01
Genetics & Heredity	828	8.43	1.14	1.02	0.92	13.18
Virology	677	17.52	0.89	1.04	0.92	16.54
Biotech. & Applied Microbiol.	633	14.96	1.35	1.02	0.92	6.89
Agriculture, Soil Science	604	41.12	1.33	0.76	0.76	8.11
Nutrition & Dietetics	587	21.84	1.31	0.97	0.80	9.31
Endocrinology & Metabolism	560	6.46	1.00	0.95	0.90	14.10
Immunology	512	4.31	0.87	1.07	0.91	17.30
Physiology	489	6.64	0.74	1.17	0.99	15.24
Horticulture	481	68.42	1.08	0.97	0.85	3.44
Neurosciences	453	2.66	0.84	1.12	1.00	16.24
Cytology & Histology	425	5.21	0.86	1.01	0.83	21.86
Multidisciplinary Sciences	328	2.91	2.09	0.89	0.77	34.79
Biophysics	319	4.54	1.05	1.09	0.98	17.32
Reproductive Systems	312	17.03	1.25	1.12	0.97	9.54

Citations per paper (C/P) was calculated from each field from the total number of citations to papers divided by the number of papers published, both taken for the years 1981–94 inclusive. RCI was calculated for each field by dividing the C/P for the group in question by the C/P for that field from all UK papers. The institutes group consisted of the following: Institute of Animal Health; Roslin Institute; Horticulture Research International; Institute for Grassland and Environmental Research; Babraham Institute; Institute of Arable Crops Research; Institute for Food Research; Silsoe Research Institute; John Innes Centre; Hannah Research Institute; Scottish Crop Research Institute; Moredun Research Institute; Rowett Research Institute; Biomathematics and Statistics Scotland; Macaulay Land Use Research Institute. The group of 15 universities was selected on the basis of their appearance in the top three in relevant fields in the recent SQ analysis⁵ with the exception that Strathclyde was replaced by Nottingham. They were the Universities of Aberdeen, Birmingham, Bristol, Cambridge, Dundee, East Anglia, Edinburgh, Lancaster, Leicester, Nottingham, Oxford, Reading, and Imperial, King's and University Colleges from London.

1. May, R. M. *Science* **275**, 793–796 (1997).

2. Bureau of Industry Economics, *Australian Science: Performance from Published Papers*, Report 96/3 (Australian Government Publishing Service, Canberra, 1966).

3. Richmond, M. H. *Sci. Public Affairs*, 2–3 (Autumn 1996).

4. *Science Citation Index* (Institute for Scientific Information, Philadelphia, 1994).

5. *Science Watch* **8** (1), 1–2 (1997).

6. Katz, J. S., Hicks, D., Sharp, M. & Martin B. R. *The changing shape of British science*, STEEP Special Report No. 3 (Science Policy Research Unit, University of Sussex, Brighton, Sussex, UK, 1995).

7. Hicks, D. & Katz, J. S. *The changing shape of British industrial research*, STEEP Special Report No. 6 (Science Policy Research Unit, University of Sussex, Brighton, Sussex, UK, 1997).

8. Krebs, J. *NERC News*, 3 (Summer 1997).

The U, J and L of bird flight

R. McNeill Alexander

Respirometric measurements of the power needed for birds to fly at various speeds have been at odds with aerodynamic theory. Direct experiments confirm theory, but loose ends remain.

Aircraft need a lot of power to fly slowly and a lot to fly fast, but fly more economically at intermediate speeds. So a graph of power against speed is U-shaped, both for fixed-wing aircraft and for helicopters (Fig. 1a). The same must be expected to be true for birds and other flying animals, yet respirometric measurements of the oxygen consumption of animals in flight have usually shown little dependence of power on speed, giving graphs that are almost flat or slightly J-shaped¹.

Oxygen consumption is not a direct measure of mechanical work, however; it indicates how much food is burned aerobically as fuel, but may not be a good indicator of work performance if muscles work with variable efficiency. Dial and his colleagues, writing on page 67 of this issue², now describe how they have made direct measurements of the work done by a flying bird. They find that, as for aircraft, much more power is needed to fly slowly than to fly at intermediate speeds.

Dial *et al.* determined the work done by the principal wing muscle of a magpie, flying in a wind tunnel. To do this, they needed to know the length changes of the muscle, and the force it exerted, at each stage of the

wing-beat cycle. They calculated the length changes from the wing movements seen in cine film. To measure the forces, they used a tiny strain gauge, implanted surgically on the crest of the humerus where the muscle attaches to it. The forces and length changes were used to calculate the work done in a wing-beat cycle, and this was multiplied by the frequency of the wing beat to obtain the mechanical power. The resulting graph of power against speed is shown in Fig. 1b. The power required for flight is high at low speeds, low at intermediate speeds and rises only slightly at high speeds. The graph is perhaps better described as L-shaped than as U-shaped.

In Fig. 1b, the graph obtained from these experiments is compared to a theoretical curve calculated for the same species using a theory devised by Pennycuik³. The theoretical curve runs parallel to the empirical one, but shows somewhat higher power requirements. This comparison is less than ideal, because Pennycuik's theory is designed to apply only to the faster of the two 'gaits' that birds commonly use in flight (the continuous vortex gait). Magpies⁴, like many other species with rounded wings⁵, seem to use the slow (vortex ring) gait even when flying at

maximum speed. Rayner⁶ has shown how power requirements can be calculated for this gait. No results from his theory have been published for magpies, but Fig. 1b includes his theoretical curve for a pigeon. This curve is U- rather than L-shaped, showing power requirements increasing steeply as the bird approaches its maximum speed. A later model⁵ predicts power requirements for the continuous vortex gait. It also produces markedly U-shaped graphs of power against speed.

Observations of gliding offer another approach. Gliding is flight powered by gravity, so a graph of the rate of loss of potential energy, against speed, is in effect a graph of power against speed. Such a graph⁷ is included in Fig. 1b. Like the theoretical graph for the pigeon, it shows power increasing steeply at high speeds.

Observations of gliding and calculations using Rayner's models^{5,6} both show power requirements increasing steeply as birds approach their maximum speeds. In contrast, Dial and colleagues² experiments on magpies show that, at the highest speed at which they could be persuaded to fly, the power requirement was well below the power they exerted in hovering.

Measurements of oxygen consumption can only tell us about metabolic power requirements in the range of speeds that can be sustained by aerobic metabolism. Hummingbirds are the only birds known to be able to hover aerobically; surprisingly, they have been found to have J-shaped curves, using oxygen no faster when hovering than when flying at moderate speeds¹. It has been suggested⁷ that this result could be an artefact of using an open-jet wind tunnel, which may have increased the power requirement at moderate speeds.

Several questions remain. Well-established aerodynamic arguments lead us to expect U-shaped graphs of power against speed. The L-shaped graph of mechanical power for magpies may be a U truncated at the right-hand side, but why is it truncated? If the power required at 14 m s^{-1} (Fig. 1b) is less than the power that the magpies exert when they hover, what is there to prevent a magpie from flying faster? The relatively flat graphs of oxygen consumption that have been obtained for various birds may be interpreted as fragments from the bottoms of U-shaped graphs, but can the J-shape of the hummingbird graph be dismissed as an artefact? □

R. McNeill Alexander is in the School of Biology, University of Leeds, Leeds LS2 9JT, UK.

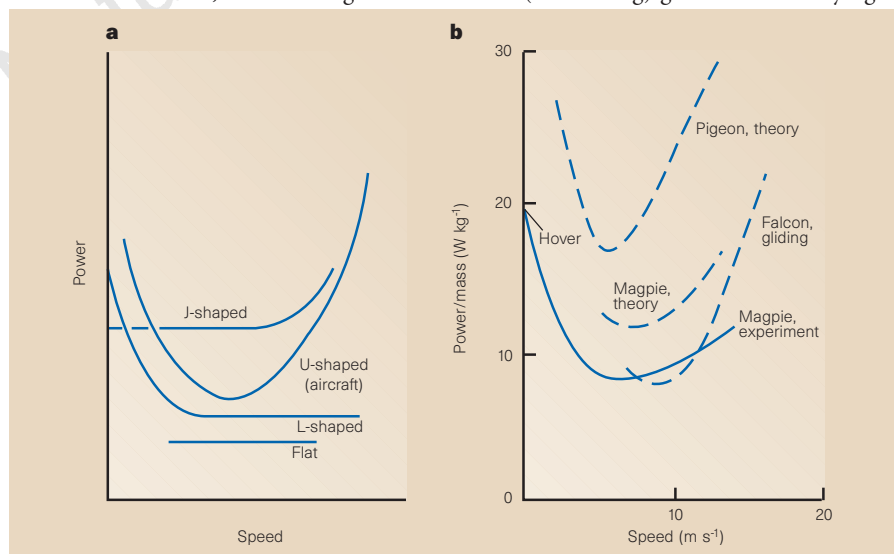


Figure 1 Power for flight at different speeds. a, Schematic graphs of power against speed, showing various possible shapes. b, Graphs of the mechanical power exerted in flight, against speed. The continuous line is the empirical graph for magpies which results from the experiments by Dial *et al.*² discussed here. Dashed lines show theoretical graphs for a magpie³ and a pigeon⁶, and a graph for a falcon calculated from its observed gliding performance⁷.

1. Ellington, C. P. *J. Exp. Biol.* **160**, 71–91 (1991).
2. Dial, K. P., Biewener, A. A., Tobalske, B. W. & Warrick, D. R. *Nature* **390**, 67–70 (1997).
3. Pennycuik, C. J. *Bird Flight Performance* (Oxford Univ. Press, 1989).
4. Tobalske, B. W. & Dial, K. P. *J. Exp. Biol.* **199**, 263–280 (1996).
5. Rayner, J. M. V. *Contemp. Math.* **141**, 351–400 (1993).
6. Rayner, J. M. V. *J. Exp. Biol.* **80**, 17–54 (1979).
7. Tucker, V. A. & Parrott, G. C. *J. Exp. Biol.* **52**, 345–367 (1970).

Structural chemistry

Oxygen bridges in molten glass

Ian Farnan

Understanding the behaviour of silicate and aluminosilicate liquids is vital to understanding the formation of igneous rocks, the explosive nature of volcanic eruptions, the viscosity of molten glass and the rate of mixing of melt batches inside industrial glass-melters. Their behaviour depends on their structure, which was thought to be well understood at a local level. But on page 60 of this issue¹, Stebbins and Xu report that a significant number of 'non-bridging' oxygen atoms are present in a glass quenched from an aluminosilicate melt — a glass that according to the conventional theory should not have them.

Silicate and aluminosilicate liquids are complex both compositionally and structurally (Fig. 1), and have high viscosities. The conventional view of their atomic-scale structure is of silicon or aluminium ions tetrahedrally coordinated by oxygen, and cross-linked into a three-dimensional network. When alkali and alkaline-earth cations are present in the melts they can modify the network structure by producing 'non-bridging oxygens' which reduce the cross-linking and decrease viscosity. But when the aluminium plus a modifier are used as a coupled substitution for silicon, it was believed that no non-bridging oxygens should be produced and a fully connected network with a higher viscosity should result.

Models of the structure of aluminosilicate liquids derive from the continuous network theory of glasses proposed by Zachariasen² and from analogy with feldspars³ (the crystalline aluminosilicate analogues of quartz). In this picture $[\text{SiO}_4]^{4-}$ tetrahedra are linked at the corners by oxygen atoms to form a three-dimensional connected structure, with all oxygens bridging between tetrahedra. A fully connected network structure can be maintained if aluminium substitutes for silicon, with the charge deficit

($\text{Si}^{4+}-\text{Al}^{3+}$) being made up by 'charge balancing' cations such as Na^+ and Ca^{2+} located close to Al^{3+} . Any excess Na^+ or Ca^{2+} will acquire non-bridging oxygens between themselves and a tetrahedral cation (Al, Si), thus interrupting the connectivity of the network or reducing the amount of cross-linking between tetrahedra. This concept has been termed 'depolymerization' and can be used to explain in general terms the very large variations in viscosity of silicate and aluminosilicate liquids with composition⁴.

This model has developed almost entirely from studying glasses quenched from the melt, and is a good starting point for characterizing the structure of the liquid at the glass transition temperature. However, it should be remembered that viscous flow and atomic transport are dynamic processes, and so we need to consider their atomic-scale rates and mechanisms. Work has begun on this for some silicate liquids using ^{29}Si NMR at high temperature^{5,6}. This has shown that the atomic-scale flow step involves an oxygen atom changing from bridging to non-bridging, or vice versa (Fig. 2). The term 'depolymerization' is thus misleading, as it is not polymeric behaviour (with fewer bridging oxygens implying shorter 'polymer' chain lengths) that decreases the viscosity of the melt, but rather the availability of more non-bridging oxygens to take part in the flow process.

But for the 'fully connected' tecto-aluminosilicate glasses, scattering and spectroscopic methods seemed to indicate a close correspondence of short- and intermediate-range structure to those of the corresponding minerals; that is, all modifiers are charge-balancing. Flow mechanisms involving non-bridging oxygens, therefore, did not seem relevant to these compositions. Now it seems that flow mechanisms similar to those of the simple silicates are probable.

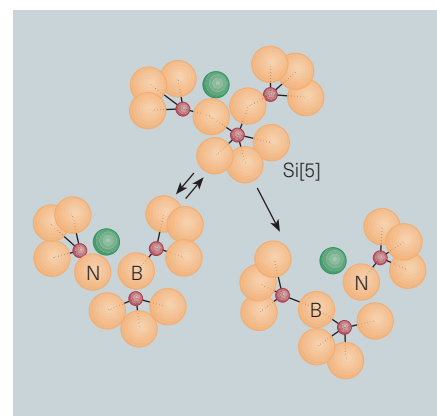


Figure 2 Cartoon of the atomic-scale flow step in a modified silicate melt. Silicons (purple), oxygens (orange), modifier (green). A non-bridging oxygen bonded to a modifier can approach a silicon and make it over-coordinated. Dissociation of the over-coordinated SiO_5 unit results in an exchange of the roles of oxygen from bridging (B) to non-bridging (N), and occurs at the right timescale to be responsible for viscosity in high-silica melts^{5,6}.

That might explain the order-of-magnitude difference in viscosity⁷ at 1,500 °C between nepheline ($\text{NaAlSi}_3\text{O}_8$) and anorthite ($\text{CaAl}_2\text{Si}_2\text{O}_8$) — in theory, two melts that have the same fully connected structure. In fact, the excess non-bridging oxygens present with calcium could be responsible for the much lower viscosity of anorthite. Indeed, detailed viscosity measurements⁸ of the system $\text{CaO}-\text{Al}_2\text{O}_3-\text{SiO}_2$ show that the viscosity is not a maximum when aluminium plus modifier exactly balance the charge of the silicon they replace, but at a slightly higher Al_2O_3 content. This may be because clusters of three aluminium atoms are present, in tetrahedral coordination⁹, which would leave excess Ca^{2+} available to form the non-bridging oxygens in the glass examined by Stebbins and Xu.

The glass industry is more than 4,000 years old, and steeped in an empirical understanding of glass and melt properties. This antiquity is probably the main reason for the industry's embracing empiricism longer than other industries, but another is the lack of any hard microscopic evidence about structure in melts and glasses (beyond simple comparisons with the local structure of crystalline analogues). The increased configurational entropy of melt and glass compared with a crystal would suggest an increase in the range of species present, but they have been difficult to identify. It appears that element-specific techniques, such as the high-resolution ^{17}O NMR used by Stebbins and Xu, are starting to get at these important differences, and glass technologists should soon be able to refine their understanding of how composition effects viscosity and atomic transport mechanisms.

The same is true for the way volcanolo-

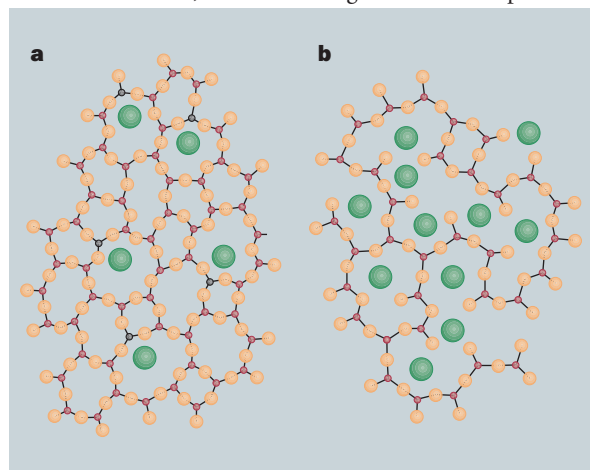


Figure 1 2D representation of the structure of an aluminosilicate network. a, In a so-called tectosilicate composition, all modifier cations (green) are conventionally thought to 'charge-balance' aluminium (grey) in the network, resulting in a structure with only bridging oxygen bonds. b, If there are too many modifier cations (or no aluminium) they disrupt the connectivity of the network by creating non-bridging oxygens. (Silicon, purple; oxygen, orange.)



Figure 3 Molten glass being snipped. The network structure makes it viscous as strong bridging-oxygen bonds need to be broken.

gists and geochemists understand the viscosity of rhyolitic magmas, which may contain more weaker, non-bridging oxygens than previously assumed. And these techniques may be even more useful in understanding explosive volcanism. As magma rises and pressure is reduced, volatiles, such as H_2O ,

come out of solution, which in a high-viscosity liquid can result in a catastrophic eruption. At present, our understanding of how a complex silicate liquid dissolves H_2O is fairly poor. H_2O can dissolve in molecular form or as OH ; the latter may bond to Si and form non-bridging oxygens. The effect on viscosity may be heavily influenced by the liquid's ability to accommodate additional non-bridging oxygens, and this will depend on the concentration already existing in the liquid.

For the sake of all these fields, we must re-examine the models we use to predict atomic transport and viscosity in these high-temperature liquids. □

Ian Farnan is in the Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK.

1. Stebbins, J. F. & Xu, Z. *Nature* **390**, 60–62 (1997).
2. Zachariasen, W. H. *J. Am. Chem. Soc.* **54**, 3841–3851 (1932).
3. Deer, W. A., Howie, R. A. & Zussmann, J. *An Introduction to the Rock Forming Minerals* 8th edn (Longman, London, 1975).
4. Mysen, B. O. *Structure and Properties of Silicate Melts* (Elsevier, Amsterdam, 1988).
5. Farnan, I. & Stebbins, J. F. *J. Am. Chem. Soc.* **112**, 32–39 (1990).
6. Farnan, I. & Stebbins, J. F. *Science* **265**, 1206–1209 (1994).
7. Urbain, G., Bottinga, Y. & Richet, P. *Geochim. Cosmochim. Acta* **46**, 1061–1072 (1982).
8. Topliss, M. J., Dingwell, D. B. & Lenci, T. *Geochim. Cosmochim. Acta* (in the press).
9. Lacey, E. D. *Acta Crystallogr.* **18**, 141–150 (1968)

Cell adhesion

How integrins are activated

Laurence A. Lasky

The integrins are a family of adhesion receptors that are involved in a diverse array of cellular processes including cell migration and growth^{1,2}. They adhere to various extracellular matrix proteins and their affinity for these ligands is altered (in response to cellular signals) by conformational changes in their ligand-binding sites^{3,4}. The α and β chains of integrin heterodimers are each composed of an extracellular region, a transmembrane segment and a small cytoplasmic domain. Integrin activation is required for cell adhesion but the molecular mechanisms involved are unclear. A paper by Fenczik *et al.*⁵ on page 81 of this issue reports that CD98, an early marker of T-cell activation, is an important regulator of integrin-mediated cell adhesion.

In previous studies, these investigators established that a limiting cellular factor is critical for adhesion involving integrins^{3,4}. They engineered Chinese hamster ovary cells to produce a chimeric protein composed of the extracellular and transmembrane portions of $\alpha_{\text{IIB}}\beta_3$ integrin and the intracellular domain of $\alpha_{6A}\beta_1$ integrin³. This chimeric protein is constitutively active, and is capable of binding ligands and transducing signals through the cytoplasmic domain. Fenczik and colleagues next engineered these cells to express another chimeric protein (Tac- β_1), composed of the

cytoplasmic region of β_1 integrin and the extracellular domain of the Tac subunit of the interleukin-2 receptor. The doubly transfected cells exhibited a dominant and concentration-dependent suppression of integrin activation, and they were no longer capable of adhesion, spreading or migration.

The suppression of adhesion in these cells might be explained by the binding of free β_1 integrin cytoplasmic domains to the putative cellular factor. Fenczik and collaborators reasoned that they might be able to identify this factor by overexpressing complementary DNAs encoding various proteins in the doubly transfected cells, and noting which protein restored cell adhesion⁵. This clever genetic complementation strategy yielded a single cDNA encoding CD98, which could overcome Tac- β_1 suppression of cell adhesion. This T-cell activation antigen is a heterodimer composed of a heavy chain with transmembrane and cytoplasmic domains, and a light chain that has yet to be characterized.

Other studies also implicate CD98 in the regulation of integrin-mediated cell adhesion. Anti-CD98 antibodies induce haematopoietic progenitor cells, which normally express CD98, to adhere to one another (homotypic adhesion), a phenomenon that is integrin dependent⁶. Furthermore,

CD98 co-precipitates with various activated β_1 integrins (cited in ref. 5). These findings suggest that the physical interaction between the cytoplasmic domains of β_1 integrin and CD98 is important in cell adhesion.

Compelling evidence also exists for a connection between CD98, integrins and virus-induced cell fusion^{7–9}. Antibodies that promote cell fusion induced by Newcastle disease virus recognize two proteins that regulate this process: α_3 integrin and the heavy chain of CD98. These antibodies also promote cell fusion induced by the gp160 envelope glycoprotein of human immunodeficiency virus, demonstrating that this phenomenon is also integrin dependent. Intracellular signalling via CD98 is clearly crucial for virus-induced cell fusion because a form of CD98 without an intracellular domain suppresses paramyxovirus-induced syncytium formation¹⁰. Moreover, such cytoplasmic deletion mutants do not inhibit Tac- β_1 suppression of cell adhesion in the Fenczik system (C. A. Fenczik, unpublished). Also, mutation of an unpaired cysteine residue in the extracellular domain of CD98 heavy chain hampers paramyxovirus-induced cell fusion. These findings are consistent with the notion that virus-induced cell fusion involves integrin-mediated adhesion events that are partly controlled by CD98.

Two simple models attempt to explain how CD98 is involved. The simple clustering model proposes that cross-linking of CD98 — by antibody, viruses or through binding of a putative ligand — brings together integrins to form a high-density complex with increased avidity for ligands because of the proximity of ligand-binding sites. This model is consistent with much of the current data, including activation by bivalent, but not monovalent antibodies. But it does not explain why CD98 that lacks a cytoplasmic domain or the unpaired cysteine in the extracellular domain disrupts virus-induced cell fusion¹⁰. Also, it is not clear how inhibitors of signal transduction could affect clustering of the CD98/integrin complex⁶.

The signalling model proposes that binding of bivalent antibodies to CD98 results in the induction of an intracellular signalling cascade that leads to a conformational change in the ligand-binding sites of integrins. This model is consistent with most of the current data, including the signalling-inhibition studies.

The findings of Fenczik *et al.*⁵ may have practical applications. Analysis of the mechanism by which CD98 regulates virus-induced cell fusion may result in the development of novel anti-viral compounds. Furthermore, if leukocytes use adhesion involving integrins and CD98 for trafficking between the bloodstream and tissues, then agents that inhibit CD98-induced upregulation of adhesion may prevent inflammation. Finally,

blocking integrin-mediated adhesion of tumour cells may be one route to preventing cancer metastasis. □

Laurence A. Lasky is in the Department of Molecular Oncology, Genentech Inc., 460 Pt San Bruno Blvd, South San Francisco, California 94080, USA.
e-mail: lal@gene.com

1. Hynes, R. *Cell* **69**, 1125 (1992).
2. Schwartz, M., Schaller, M. & Ginsberg, M. H. *Annu. Rev. Cell Dev. Biol.* **11**, 549–599 (1995).

3. Chen, Y. *et al. J. Biol. Chem.* **269**, 18307–18310 (1994).
4. LaFlamme, S., Thomas, L., Yamada, S. & Yamada, K. *J. Cell Biol.* **126**, 1287–1298 (1994).
5. Fenczik, C. A., Sethi, T., Ramos, J. W., Hughes, P. E. & Ginsberg, M. H. *Nature* **390**, 81–85 (1997).
6. Warren, A., Patel, K., McConkey, D. & Palacios, R. *Blood* **87**, 3676–3687 (1996).
7. Ohta, H. *et al. EMBO J.* **13**, 2044–2055 (1994).
8. Ohgimoto, S. *et al. J. Immunol.* **155**, 3582–3592 (1995).
9. Ohgimoto, S. *et al. J. Gen. Virol.* **77**, 2747–2756 (1996).
10. Okamoto, K. *et al. J. Gen. Virol.* **78**, 83–89 (1997).

Photonics

A photon in solitary confinement

Yoshihisa Yamamoto

In many ways, photons and electrons have opposite natures. A photon has no electrical charge and so does not directly interact with other photons via the ordinary Coulomb force; and a group of photons obey Bose–Einstein statistics. Electrons, on the other hand, interact strongly and are inherently nonlinear, because they are charged and obey Fermi–Dirac statistics. The non-interacting aspect of photons is very useful for transmitting information without signal distortion, but is ill-suited for information processing, which requires nonlinear interaction between particles. So the current belief is that communication links should be based on non-interacting photons, but computer logic and memory should be constructed by interacting electrons.

But photons do interact with each other in a nonlinear medium. Until now, however, a trade-off between nonlinearity and loss has prevented optics from going beyond simple communication and sensor applications. A new method called ‘atomic dark resonance’, proposed by Imamoglu and colleagues¹, may change that. As in earlier schemes, they propose a nonlinear cavity to make photons interact strongly, but they have discovered a way to avoid the usual absorption losses. Thus they can increase the available nonlinearity by almost ten orders of magnitude for a given light intensity and loss. Because of this huge increase, it is in principle possible to control and switch an optical cavity with a single photon.

A photon–photon interaction is only possible when the refractive index of a medium depends on the light intensity (a property called optical Kerr nonlinearity). When the photon energy is close to an atomic transition energy (within about one transition linewidth), the Kerr nonlinearity is resonantly enhanced and the interactions are stronger; but absorption simultaneously increases, which makes this type of resonant nonlinearity impractical for most applications. The new scheme uses a carefully chosen coherent superposition of the two atomic ground states. In steady state, almost all the atoms are in the stable, lower-energy ground state, with a small virtual population in the metastable,

higher-energy ground state which is created by a coherent driving field, and thus it is possible to eliminate absorption using destructive interference between the transitions from the two ground states to a common excited state. This phenomenon is known as electromagnetically induced transparency, or atomic dark resonance². Because one-photon loss is negligible, the photon energy can approach the atomic transition energy much closer than the linewidth limit and so a giant Kerr nonlinearity is possible.

Two other methods are being pursued elsewhere. One is the search for a medium with a large dipole matrix element between the ground and excited states, which means having highly polarizable atoms or molecules. The Kerr nonlinearity is proportional to the fourth power of the matrix element, whereas the one-photon absorption loss is proportional to the square of the matrix element, so a larger matrix element promises higher Kerr nonlinearity for a given linear loss. Efforts along this line have not been very successful, even though there is no fundamental principle which forbids the existence of such an ideal nonlinear medium. Another approach, similar to the one by Imamoglu and colleagues¹, is self-induced transparency³. In this special scheme, the leading edge of an optical pulse virtually excites all the atoms in the ground state to the excited state, but the trailing edge of the pulse induces the stimulated emission of photons and brings all the atoms back to the ground state. Thus, the optical pulse propagates without absorption even though the optical field resonantly couples with the atoms. But to avoid one-photon loss, a special condition must be satisfied for the pulse intensity and width, whereas no such constraint exists in electromagnetically induced transparency.

One consequence of this giant Kerr nonlinearity is that an optical cavity can be controlled by single photons. When the cavity is empty, the incident photon energy satisfies the resonance condition and the first photon goes into the cavity via resonant optical tunnelling. However, once a single photon is trapped, the resonance is shifted owing to Kerr nonlinearity, so subsequent photons



100 YEARS AGO

An instructive fisheries exhibition, arranged to illustrate the fishing industries and the application of science to agriculture, was opened in the Zoological Museum of the University College, Liverpool, on Friday last. ... As Prof. Herdman pointed out at the opening ceremony of the exhibition, fishes are animals, their food is composed of animals, and their enemies are animals. In all the operations of their life, their feeding and breeding, and so on, they are subject to the same biological laws which regulate the lives of all animals in the sea. The investigation of fishery questions is applied biology, and if our fisheries are to be benefited they must be treated in a scientific manner. We do not trust to unaided nature for our supplies of breadstuffs and beef. Why then should we trust to nature for fish? As there is an agriculture of the land, so there must be an aquiculture of the sea. Fishermen must in the future be farmers of the sea-shore, not hunters as they have been in the past. From *Nature* 4 November 1897.

50 YEARS AGO

During the present century, a new ‘slave-trade’ has arisen. Individuals are torn from their native element and are made to perform functions entirely foreign to them; immense coercion is applied to make them perform these functions; and when they have performed them they are carelessly discarded. Since the existence of these individuals – electrons – was first recognized fifty years ago, man has transformed out of all recognition his power and control over them. We are now not sure of the exact nature of the electron. Is it a particle or a wave packet? But whatever it is, wave-mechanics leads to no suggestion that the electron has a soul; and so man can, for the present, continue his slave-trade without any qualms of conscience. The exhibition arranged by the North-West Branch of the Institution of Electronics and held at the College of Technology, Manchester, during July 22–23, might justly be described as a tour of the torture chambers of the electrons, the scale of tortures being finely graded from the mere interchange of habitat, as in accumulators, to the dragging out into free space and violent oscillation that occurs in high-frequency valves. From *Nature* 8 November 1947

cannot enter until the trapped photon escapes. Imamoglu and colleagues call this a "photon (Kerr) blockade", by analogy with the electron (Coulomb) blockade⁴. In an electron blockade, the incident electron energy is just right to resonantly tunnel into the central island of a double-barrier tunnel diode. Once it is trapped, the resonant energy is shifted due to Coulomb repulsion, and electrons arriving later can't get in. This effect underlies various single-electron devices such as transistors, turnstiles and pumps⁴.

Similarly, the photon blockade will probably be used in devices such as single-photon turnstiles⁵ and single-photon quantum logic gates^{6,7}, which may be important for quantum cryptography and computation. Experimental demonstrations of the proposed

giant Kerr nonlinearity and photon blockade are yet to be seen, but the work of Imamoglu *et al.* is an important step towards optical quantum information technologies. □

Yoshihisa Yamamoto is in the ERATO Quantum Fluctuation Project, Edward L. Ginzton Laboratory, Stanford University, Stanford, California 94305-4085, USA; and NTT Basic Research Laboratories, Atsugishi, Kanagawa, Japan.

1. Imamoglu, A., Schmidt, H., Woods, G. & Deutsch, M. *Phys. Rev. Lett.* **79**, 1467–1470 (1997).
2. Harris, S. E. *Physics Today* **50**, 36–42 (1997).
3. Watanabe, K., Nakano, H., Honold, A. & Yamamoto, Y. *Phys. Rev. Lett.* **62**, 2257–2260 (1989).
4. Grabert, H. & Devoret, M. H. *Single Charge Tunneling* (Plenum, New York, 1992).
5. Imamoglu, A. & Yamamoto, Y. *Phys. Rev. Lett.* **72**, 210–213 (1994).
6. Brune, M. *et al.* *Phys. Rev. Lett.* **77**, 4487–4490 (1996).
7. Turchette, O. A. *et al.* *Phys. Rev. Lett.* **75**, 4710 (1995).

Ageing

New mice for old questions

George M. Martin and I. Saira Mian

In 1733, Alexander Pope¹ wrote that man was "Created half to rise, and half to fall". In other words, "First we ripen, then we rot"! Pope, it seems, was 250 years ahead of his time in anticipating the modern evolutionary theory of ageing^{2,3}. This theory states that organisms in age-structured populations senesce because of the weakening of the force of natural selection as they grow older, which allows certain gene actions to reach deleterious levels of expression. In human populations, common examples include cataracts, Alzheimer's disease and other dementias, type 2 diabetes mellitus, osteoporosis (brittle bones), many types of cancer and several types of arteriosclerosis (hardening of the arteries). On page 45 of this issue, Kuro-o and colleagues⁴ describe a single mutation in a mouse gene that, when homozygous, appears to accelerate several types of pathology associated with late-life disorders in humans (but not in mice).

Like hundreds of labs that introduce bits of genetic material into mice in the hopes of getting expression of such foreign genetic information, these investigators inadvertently made a random 'hit' that crippled one of the host's normal genes. Unlike many such labs, however, they decided to alter course and to investigate what seemed to be a model of accelerated ageing in man. We can all be happy that they made this decision, as the results are of considerable interest, although not fully explained and of questionable relevance to mechanisms of human senescence. A special merit of their study is that they re-introduced a normal version of the gene into the mutant animals and essentially normalized the phenotype, thus satisfying a sort of genetic equivalent of Koch's postulates.

They named their gene *klotho* (*kl*) after one of the Fates of Greek mythology (Pope, a

translator of Homer, would have approved). The gene encodes a cell-surface protein of 1,014 amino acids with a putative signal sequence at the amino terminus and a single transmembrane helix near the carboxy terminus. The extracellular domain is composed of two internal repeats (KL1, KL2), each of which has 20–40% sequence identity to the β -glucosidases of bacteria and plants as well as to mammalian lactase glycosylceramidase (Fig. 1).

The sequence similarity between Klotho

and mammalian lactase glycosylceramidase suggests a role for Klotho in sphingolipid metabolism. From the diverse phenotypes of *kl* mutant mice, it seems that extracellular insults or agents such as hormones could activate Klotho, which might then act on membrane glycolipids and gangliosides to release ceramide — a molecule involved in cell replicative senescence, apoptosis and cell-cycle arrest. Given the predicted extracellular location of KL1 and KL2, Klotho probably acts on outer-leaflet membrane lipids. So, for ceramide generated by Klotho to initiate an intracellular response, it would need to diffuse through the membrane into the cytoplasm or alternatively to act extracellularly. The latter would require the existence of molecules on the cell surface and/or in the extracellular matrix that could interact with ceramide. Klotho may perhaps be involved (either directly or indirectly) with remodelling of the extracellular matrix.

Although the authors argue that their *kl* mutants develop normally, it is clear that the appearance of symptoms as early as 3–4 weeks after birth means that their mice have not fully 'ripened'. (The equivalent human age is six years.) But for each locus identified on the basis of a strong mutational effect, there could exist 'leaky' mutants and, perhaps of more general importance to the pathobiology of ageing, polymorphic variants. Such genetic variants could result in significant modulation of phenotypes that emerge after the reproductive phase of life is finished. At the level of precise phenotypic analysis, however, there are important

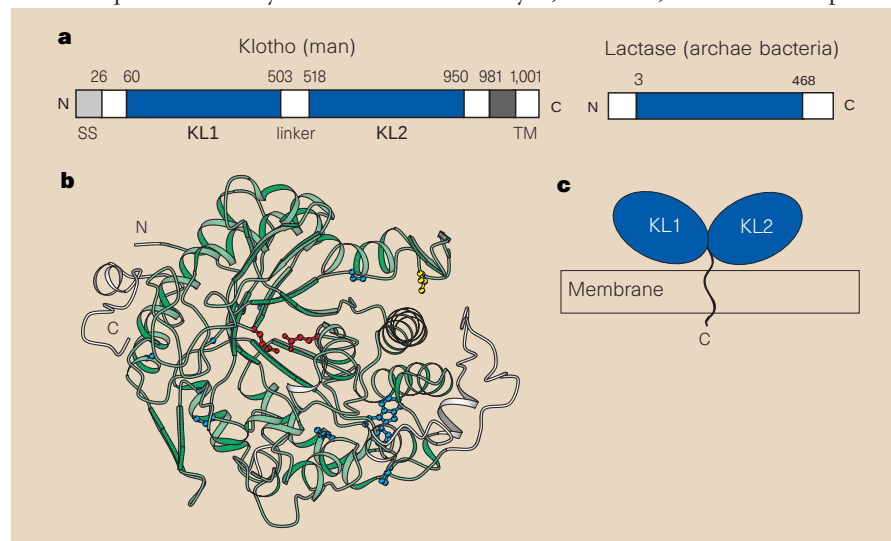


Figure 1 a, Comparison of human Klotho and archaeal lactase, proteins that contain β -glycosidase domains (labelled KL1 and KL2 in Klotho). b, The three-dimensional structure of archaeal lactase is shown as a ribbon diagram (conserved residues are turquoise, glutamic-acid residues in the active site are red, and the regions present in KL1 and KL2 but absent in lactase are yellow). Although KL1 and KL2 domains in Klotho are likely to have tertiary structures similar to that of lactase, subtle differences in the active site and other residues may exist. c, The spatial proximity of the amino and carboxy termini of lactase and the relatively short linker region between KL1 and KL2 suggest that Klotho may have a butterfly shape. (The common features of these proteins were compared using a hidden Markov model^{7–9}. The model is at URL <http://genome.wustl.edu/eddy/hmm.html>; software is at URL <http://www.cse.ucsc.edu/research/compbio/sam.html>)

caveats that lead one to question the relevance of the *kl* mutation to normal human ageing. For example, osteopenia (decreased bone mass) in mutant mice cannot be immediately equated with human osteoporosis. Moreover, the conclusion that there is atrophy of anterior pituitary cells and associated deficiency of growth hormone requires quantitative morphological and biochemical assessments. In any case, there is evidence that mice that are genetically deficient in growth hormone live longer than controls⁵.

What is known about β -glucosidase mutations in man? Perhaps the closest (but quite imperfect) match to the *kl* mutant phenotype is Gaucher's disease⁶. Patients are homozygous for a rare point mutation at the (lysosomal) β -glucocerebrosidase locus on the long arm of human chromosome 1. Klotho appears to have a different subcellular localization (plasma membrane). Unlike the classical forms of Gaucher's disease, *kl* mutant mice have only sparse numbers of cells engorged with unmetabolized substrate. These mice also show progressive cardiovascular calcification, hydrocephalus, selective neuronal loss, opaque corneas, deafness and deformed toes. One wonders if the abnormal strides of *kl* mutant mice are suggestive of subtle foot deformities rather than of Parkinson's disease, as the authors

suggest. The strong *kl* expression in the choroid plexus hints that some degree of hydrocephalus might have been missed in these mutant animals.

The authors have mapped the human *kl* gene to the long arm of chromosome 13. It is therefore possible that a yet-to-be-described mutation at this locus could give a phenotype in man that is more closely related to the lesions seen in mice homozygous for the *kl* mutation. If so, some variants at this locus may satisfy Pope's conception of proper ageing. □

George M. Martin is in the Department of Pathology, University of Washington, Seattle, Washington 98195, USA.

e-mail: gmmartin@u.washington.edu

I. Saira Mian is in the Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA.

1. Pope, A. *An Essay on Man* (ed. Brady, F.) (Bobbs-Merrill, Indianapolis, IN, 1965).
2. Medawar, P. B. (ed.) in *The Uniqueness of the Individual* 44–70 (Methuen, London, 1957).
3. Williams, G. C. *Evolution* 11, 398–411 (1957).
4. Kuro-o, M. et al. *Nature* 390, 45–51 (1997).
5. Brown-Borg, H. M., Borg, K. E., Meliska, C. J. & Bartke, A. *Nature* 384, 33 (1996).
6. Abrahamov, A. et al. *Lancet* 346, 1000–1003 (1995).
7. Krogh, A., Brown, M., Mian, I. S., Sjolander, K. & Haussler, D. *J. Mol. Biol.* 235, 1501–1531 (1994).
8. Baldi, P., Chauvin, Y., Hunkapiller, T. & McClure, M. A. *Proc. Natl Acad. Sci. USA* 91, 1059–1063 (1994).
9. Eddy, S. R. *Curr. Opin. Struct. Biol.* 6, 361–365 (1996).

Evolutionary biology

Sex and synergism

Emily J. Lyons

“In short, it seems certain — and here we must simply grit our teeth and swallow — that it is impossible to understand the function of sex without both a firm grasp of the rival hypotheses, their assumptions, logical development and consequences, and also a reasonably comprehensive knowledge of the nature, occurrence and correlates of sexual systems in nature.” Graham Bell¹ wrote this 15 years ago. His comments illustrate, in a nutshell, the reasons behind a conference held late this summer to examine the enduring mystery of the prevalence of sexual reproduction*.

The event was marked by the first synthesis of ideas in the field since 1988 (ref. 2), and by a shift from theoretical approaches to empirical tests of both the function of sex in natural populations and of the assumptions underlying the various explanations for it. The meeting culminated primarily in consensus and a clear agenda for future research. However, some key theories for the maintenance of sex hinge on the rate at which deleterious mutations accumulate in the genome of an organism, and there was much

debate about the measurement of this rate.

How is it that there are so many sexual organisms when a female could have twice as many grandchildren if she reproduced asexually? Over a century ago, Weismann proposed that the significance of sexual reproduction is “the creation of hereditary individual characters to form the material upon which natural selection may work”³. Contemporary workers agree with Weismann, for all of the current, plausible theories hold that sex is advantageous because it creates variation, the raw material for evolution by natural selection, and results in either the assembly of mutations that confer an advantage or the clearance of those that do not (L. Hurst, Univ. Bath; G. Bell, McGill Univ.).

The usefulness of one model of mutation clearance, the ‘mutational deterministic’ hypothesis (A. Kondrashov, Cornell Univ.) received close attention. According to this view, if deleterious mutations interact synergistically (that is, each successive mutation has progressively larger effects on the decline in an organism's fitness), then asexual populations will have a higher mutation load than sexual populations at equilibrium — the point at which the accumulation of mutations is equal to their removal by natural

selection. If the rate at which deleterious mutations occur is greater than about one per genome per generation, then this load should result in an advantage for sexuality^{4,5}.

It was originally thought that the beauty of this model lay in its falsifiability; clearly, if the mutation rate in a sexual population is less than one per genome per generation, then the model cannot explain why sex is maintained. The difficulty arises with the fact that one cannot simply look at an organism and know its deleterious mutation rate — at best, that rate can be measured only indirectly.

The issue is further complicated by disagreement about the mutational process in general. For example, is the deleterious mutation rate roughly constant across the genome of an organism (A. Kondrashov; B. Tautz, Univ. Munich)? Or are some genes in regions of the genome with a very low mutation rate whereas other regions are prone to higher rates (Hurst)?

The measurement of the rate, constant or otherwise, is similarly contentious. Typically, the measurement is made by calculating the proportion of the genome that is constrained by selection, and the mutation rate per base pair in that constrained region. Tautz estimates the size of the region using the number of amino acids coded for, assuming that the greater the number of the latter, the more highly constrained the region. Kondrashov, however, believes that estimates of this sort may be highly conservative. Not surprisingly, these disparate views on the mutation process and how best to measure mutation rates led to estimates that ranged from 0.06 to 10 per genome per generation, making the mutational deterministic model less easily falsifiable than originally thought.

On the other side of the coin (mutations that are advantageous in themselves or when assembled together), the Red Queen hypothesis postulates that host–parasite coevolution is the critical force. Sex and recombination are favoured because they allow for the production of new, rare genotypes that are expected to have a greater chance of escaping parasites or, in the case of parasites, of infecting new hosts⁶. The operation of this model is perhaps best seen in natural populations of a New Zealand freshwater snail, *Potamopyrgus antipodarum*, where the frequency of sex is highest in lakes and habitats within lakes where parasites are most prevalent (C. Lively, Indiana Univ.). Parasites can play the same game, only in reverse, as evidenced in their increased sexual reproduction when faced with host immunity. Specifically, parasitic nematode larvae from rat hosts that have acquired immune protection are more likely to develop into sexual adults than those from unprotected hosts (A. Read, Univ. Edinburgh).

Nonetheless, there are systems where sex

* Summer School on the Evolution and Maintenance of Sex, Max Planck Institute for Behavioural Physiology, Seewiesen, Germany, 10–17 August 1997.

is maintained in the absence of parasites. For example, in laboratory studies of the unicellular green alga *Chlamydomonas reinhardtii*, sex can serve both for mutation clearance and mutation assembly although no parasites are involved (Bell). In consequence, a more pluralistic view of the maintenance of sex may well be appropriate — it seems reasonable that sex might be maintained in different populations for different reasons, and that the answer to the problem of sex may be relative. “Maybe” is the operative phrase, for we don’t know nearly enough to reach a solid conclusion. The task for the future is to

achieve that “reasonably comprehensive knowledge of the nature, occurrence and correlates of sexual systems in nature”¹. □

Emily J. Lyons is in the Department of Biology, Indiana University, Bloomington, Indiana 47405, USA.

e-mail: elyons@sunflower.bio.indiana.edu

1. Bell, G. *The Masterpiece of Nature: The Evolution and Genetics of Sexuality* (Univ. California Press, Berkeley, 1982).
2. Michod, R. E. & Levin, B. R. *The Evolution of Sex* (Sinauer, Sunderland, MA, 1988).
3. Weismann, A. *Essays Upon Heredity* (Clarendon, Oxford, 1889).
4. Kondrashov, A. S. *Genet. Res.* **40**, 325–332 (1982).
5. Kondrashov, A. S. *Nature* **336**, 435–440 (1988).
6. Hamilton, W. D. *Oikos* **35**, 282–290 (1980).

Palaeontology

Fins made for walking

Oleg A. Lebedev

One of the most enigmatic issues in vertebrate evolution is the fish–tetrapod transition — that period of evolutionary history when, some 365 million years ago, four-legged land animals first appeared in the fossil record. The transition was accompanied by radical innovation in limb structure such as the appearance of digits, and later (although we don’t know exactly when) adaptations to habitual air breathing.

What were the causes and mechanisms of

these changes? Such questions were matters of speculation until recently. Limbs capable of supporting an animal on land have been thought of as adaptations to terrestriality adopted by osteolepiform fish, the group from which tetrapods are thought to have arisen. But almost 40 years ago Romer¹ wondered whether the fin–limb transition took place in water, independently of walking as such. A series of papers from Jennifer A. Clack and Michael I. Coates on the most primitive known tetrapod, the amphibian

Acanthostega gunnari from the Upper Devonian of East Greenland, suggests that Romer was right. In a definitive study, published in *The Transactions of the Royal Society of Edinburgh*², Coates has gone further and set out the postcranial evidence and its implications for phylogeny and evolution.

The retention of numerous fish-like characters in the skull^{3–5} and postcranial skeleton^{2,5,6} shows how primitive *Acanthostega* is in relation to other contemporary and later tetrapods. For example, it has a well-developed postbranchial lamina, which forms the support of the rear wall of the branchial chamber — something to be expected in a fish, but a surprising find in a tetrapod. The animal also retains an anocleithrum. In fish, this bone connects the horizontal branch of the pectoral girdle (attached to the back of the skull) with the vertical branch (what in modern tetrapods we would now call the shoulder girdle), but it is not a bone one associates with tetrapods. And yet, in contrast to fishes, *Acanthostega* lacks the opercular series and, correspondingly, the horizontal branch of the girdle, the anocleithrum remaining as a relic of the past connection to the skull. This reduction may have resulted from transformation of the respiratory system. It made possible significant enlargement of the orobranchial cavity for the purpose of pumping air with the help of the vertical movements of the floor of this

Remote sensing

Night lights

New satellite images of the Earth at night give the most complete view of global human settlement yet obtained. The Operational Linescan System (OLS) detects visible and near-infrared emissions resulting from human activity. These maps not only reveal a fascinating view of worldwide ‘light pollution’, but they also serve as indirect tracers of international greenhouse-gas production.

The OLS sensors are part of the Defense Meteorological Satellite Program, and have been operational for two decades. But previous images suffered limited areal coverage and were unable to filter out misleading signals, such as temporary light from fires.

Using new techniques to process digitally archived data, C. D. Elvidge and colleagues (*Glob. Change Biol.* **3**, 387–395; 1997) have circumvented these problems. By analysing a series of images acquired at different times during 1994–95, they have built up maps of the permanent light and heat emission from the Americas, Asia and Europe (shown here). Moreover, by selecting observations acquired when there was little moonlight (which can bounce off



clouds and produce spurious signals), they can achieve optimal sensitivity to faint, small sources at ground level.

Variations in the density of the emitted light are striking but not surprising — certainly, the major cities cannot hide. However, the maps also facilitate comparisons between the illuminated area of each country, and statistics such as

population and carbon emission.

The OLS data suggest that, allowing for certain biasing factors, these statistics are directly correlated with the area of lit regions. So the results may provide a way of gauging different national contributions to greenhouse-gas production, and the wider implications for climatic change.

Karen Southwell

cavity and, at the same time, released the head from rigid connection with the trunk to facilitate horizontal and vertical movements.

The sacral part of *Acanthostega's* vertebral column is only slightly differentiated from the fish-like state, and there are almost no specialized vertebrae that bear sacral ribs for the attachment of the pelvis to the vertebral column. In fish, the pelvis is not connected to the vertebral column, which remains undifferentiated. In tetrapods, walking depends on an effective mechanical connection between the axial skeleton and the ground, so a link between the pelvis and the vertebral column, through the sacrum, becomes an advantage. Just like a fish, *Acanthostega* has a tail-fin, supported by bony rods of dermal origin called lepidotrichia. In contrast to all known lobe-finned fish, however, the lepidotrichia are unbranched and unsegmented.

Coates² shows that fish-like characters in *Acanthostega* tend to be concentrated in the posterior parts of the animal, suggesting that the initial stages in the fish–tetrapod transformation concerned the head and anterior parts of the trunk — especially those parts, such as the opercular region and forelimb, that concerned ventilation and locomotion.

As expected, limbs and their girdles were especially strongly affected during the fish–tetrapod transition, but many aspects of the replacement of fins by digits remain unclear. Coates suggests that shorter appendages were more suitable than long paddle-like fins rimmed with lepidotrichia for navigating shallow, vegetation-filled basins^{2,5}. A change of function, from exclusively paddling, to a combination of paddling and 'walking' along the bottom, required internal bracing in the form of an enlarged pelvic girdle and elements in the pectoral girdle such as the scapulocoracoid.

At first, the forelimbs were used passively, as props to lift the animal clear of the bottom. Walking ability developed later. The primitive construction of the *Acanthostega* forelimb is revealed in the fish-like proportions of the radius and ulna compared with those of later forms, although some adaptation to tetrapod-like locomotion is revealed by the shape of the humerus. But the wrist bones were not ossified and the ankle joint was absent, although in the contemporaneous tetrapod *Tulerpeton* from central Russia this structure is already evident⁷. Clearly, the evolution of ankle joints and wrist joints lagged behind the formation of digits.

But what of those digits? The traditional textbook view^{8,9} sees tetrapods as primitively pentadactyl (with five digits on each limb), with some attempt to trace each digit back to the bony elements in osteolepiform fin-skeletons. This view stood until 1984, when I published¹⁰ an account of the hexadactylous (six-digitated) *Tulerpeton curtum*. After that, polydactyly became a matter of interest

among students of tetrapod origins. *Acanthostega* manifests an even more primitive octodactylous (eight-digitated) condition, and polydactyly is now the rule in all Devonian tetrapods for which digits are known¹¹. That polydactyly is primitive fits in well with embryological data^{12,13} showing that digit number is not fixed, but relates to the extent of the apical ectodermal ridge zone during ontogeny. The larger the ridge, the more digits are likely to form.

'Transitional' fossils such as *Acanthostega* allow us to work out not only the changes necessary for life on land, but the order in which they occurred. For example, Coates² shows that the formation of the tetrapod limb went well ahead of changes in the respiratory system: *Acanthostega* had digits, but it also had fully functional internal gills. But modern lungfish — the closest living relatives of the tetrapods — show that the order in which such changes occurred need not have been inevitable, and that there is more than one way to adapt to life on land. In the African lungfish *Protopterus*, for example, the pectoral fin skeleton remains primitive and, more, becomes greatly reduced, but the lungs are effective. As in *Acanthostega*, the opercular skeleton becomes lost but functional internal gills are retained.

Acanthostega is more primitive in many ways than *Tulerpeton* and another contemporary amphibian, *Ichthyostega*. This implies that not only was it more adapted to life in water, but that it originated directly from water-dwelling ancestors, and indeed may never have moved onto land. Despite this, the mixture of fish-like and tetrapod-like features in *Acanthostega* shows that it is a genuine transitional form which, as such, sheds much light on the transition from water to land. The revival of interest in the fish–tetrapod transition, fuelled by fossil discoveries over the past few years, is continuing, and Coates's monograph on *Acanthostega* will remain the standard work on the subject for many years to come. □

Oleg A. Lebedev is in the Palaeontological Institute of the Russian Academy of Sciences, 123 Profsoyuznaya St, Moscow 117647, Russia.
e-mail: olebed@paleo.msk.su

1. Romer, A. S. *Evolution* 12, 365–369 (1958).
2. Coates, M. I. *Trans. R. Soc. Edinb. Earth Sci.* 87, 363–421 (1996).
3. Coates, M. I. & Clack, J. A. *Nature* 352, 234–236 (1991).
4. Clack, J. A. *Medd. Grönland Geosci.* 39, 1–24 (1994).
5. Clack, J. A. & Coates, M. I. *Bull. Mus. Natl Hist. Nat., Paris* 17 (C), 359–372 (1995).
6. Coates, M. I. & Clack, J. A. *Bull. Mus. Natl Hist. Nat., Paris* 17 (C), 373–388 (1995).
7. Lebedev, O. A. & Coates, M. I. *Zool. J. Linn. Soc.* 114, 307–348 (1995).
8. Jarvik, E. *Basic Structure and Evolution of Vertebrates* (Academic, London, 1980).
9. Carroll, R. L. *Vertebrate Paleontology and Evolution* (Freeman, New York, 1988).
10. Lebedev, O. A. *Dokl. Akad. Nauk* 278, 1470–1473 (1984).
11. Coates, M. I. & Clack, J. A. *Nature* 347, 66–69 (1990).
12. Shubin, N. H. & Alberch, P. *Evol. Biol.* 20, 319–387 (1986).
13. Hinchliffe, J. R. in *Complex Organismal Functions: Integration and Evolution in Vertebrates* (eds Wake, D. B. & Roth, G.) 171–189 (Wiley, New York, 1989).

Daedalus

A radical steam-trap

Electrolysis is the neatest and most controllable of chemical reactions. Its applications are sadly limited by the problem of the electrodes. They are always attacked by discharging ions. Only in a few cases is this tolerable, or chemically useful. Daedalus now has a way out.

He points out that a smooth, red-hot heating element boils water very inefficiently. It becomes completely covered with a thin blanket of steam, through which heat leaks only slowly. Imagine, he says, such an element used as a electrode in an electrochemical cell. Its hot surface would emit electrons readily, like a valve cathode, and so should the hot water surface. Electrons travel quite well through water vapour (the special scanning electron microscope for wet specimens exploits this effect), though to aid their passage the system should be maintained at as low a pressure as possible. Two such electrodes would form an electrolytic cell. Most of its applied voltage would be dropped across the thin steam blankets around the electrodes, but this at least would help to keep them hot. Only a fraction of the voltage would go into electrochemistry; but that fraction would perform wonderful new reactions.

An ion discharged at a steam electrode will, as usual, be converted to the corresponding radical — often ferociously reactive and unstable. But deprived of the normal easy option of attacking the electrode, this radical will vent its energy on the molecules around it. Hydrogen ions will go to that energetic species, monatomic hydrogen. Metal ions will go to isolated metal atoms, hitherto available for chemistry only in the vapour phase. Azide will go to N₃, chlorate to ClO₃, hydroxide to OH — the most ruthlessly reactive of all radicals. All these will attack whatever target molecules the chemists have added to the solution; or failing that, the water itself. The results should be highly entertaining, and even useful. The vast new array of hitherto ignored or unknown radicals will make all sorts of splendid new syntheses possible.

But as a first step, Daedalus wants to use steam electrolysis to purify water. A few ferocious radicals should effortlessly degrade the few parts per billion of dioxin and halocarbons which make environmentalists faint with fear. All the nasties will be instantly mineralized, as the jargon has it. The most dubious piped product will be neatly converted into a chic mineral water.

David Jones

Galileo at Jupiter — meetings with remarkable moons

William B. McKinnon

The four large moons of Jupiter form the most coherently organized planetary system known. Over the past two years, the Galileo spacecraft has illuminated both the interconnections between these worlds and the uniqueness of each, challenging theories of moon formation and evolution.

In January of 1610, Galileo Galilei turned his newly fashioned telescope upon Jupiter, finding "... four stars that wander around Jupiter as does the moon around the earth..."¹. These were not stars, of course, but satellites of the great planet; and it is fitting that four centuries later a sophisticated robot named in his honour orbits Jupiter as well. Its mission: to seek out new knowledge of the planet, and by means of nearly a dozen close passes, to explore in detail the four Galilean moons — Io, Europa, Ganymede and Callisto.

Most of the basic parameters of Jupiter's

satellite system have been known since the Voyager flybys of 1979 (see Table 1, overleaf). The inner two Galilean moons (Fig. 1), Io and Europa, are about the size of our Moon, and composed mostly or entirely of rock and metal. Ganymede and Callisto are larger, equal to or greater than Mercury in size, and roughly half ice by mass. This division is reminiscent of the Solar System as a whole, where the rock- and metal-rich inner planets are distinct from the much larger gas- and ice-rich outer planets. Jupiter's satellite system is, however, more 'systematic', in that

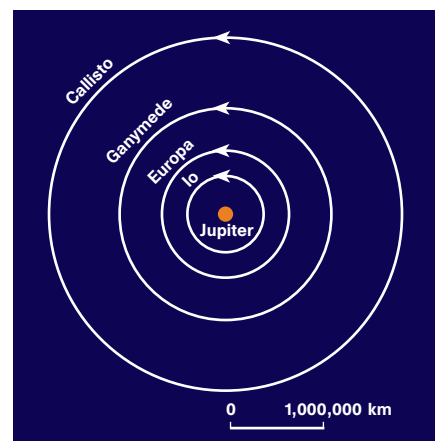


Figure 1 The orbits of the four large satellites of Jupiter. Many much smaller moons are also known.

most properties vary continuously with distance from Jupiter. All of the bodies are essentially solid, like the Earth, but vary by water-ice fraction. Io is ice-free, Europa has a surface shell of ice (and possibly water), and although Ganymede and Callisto are both ice-rich, outermost Callisto has more.

This compositional gradient has curious geological parallels. Io is extremely geologically active (an understatement, actually), Europa seems also to be active, though on a more modest scale, and Ganymede has undergone bouts of activity in its geological past. Only Callisto appears to have refrained from revealing any hints of internal activity. Correspondingly, Callisto's surface is very heavily cratered, Ganymede is heavily cratered in parts but more or less resembles the Moon in terms of the range of crater densities, Europa is very lightly cratered, and Io lacks any detected impact craters at all, so far. This is all the more remarkable given that Jupiter's gravity concentrates the flux of comets and asteroids passing near it, which substantially increases the bombardment rate of the inner moons with respect to the outer ones².

Jupiter is principally responsible for these patterns. Early in its history, Jupiter was larger and more luminous than it is today (ref. 3, for example). As proto-Jupiter acquired its vast gaseous envelope from the solar nebula, excess angular momentum forced some of the material into a flattened disk orbiting in the planet's equatorial plane, forming a planetary counterpart to the much larger solar nebula. The condensable material in the protojovian nebula, once gathered together (accreted) by gravity, created the satellite system. So it is by birth as well as by organization that the Galilean satellites, along with Jupiter, can be called a miniature planetary system.

Jupiter's early luminosity ensured that the inner regions of the protojovian nebula remained relatively warm. Rock and metal, but not water ice, should have condensed in

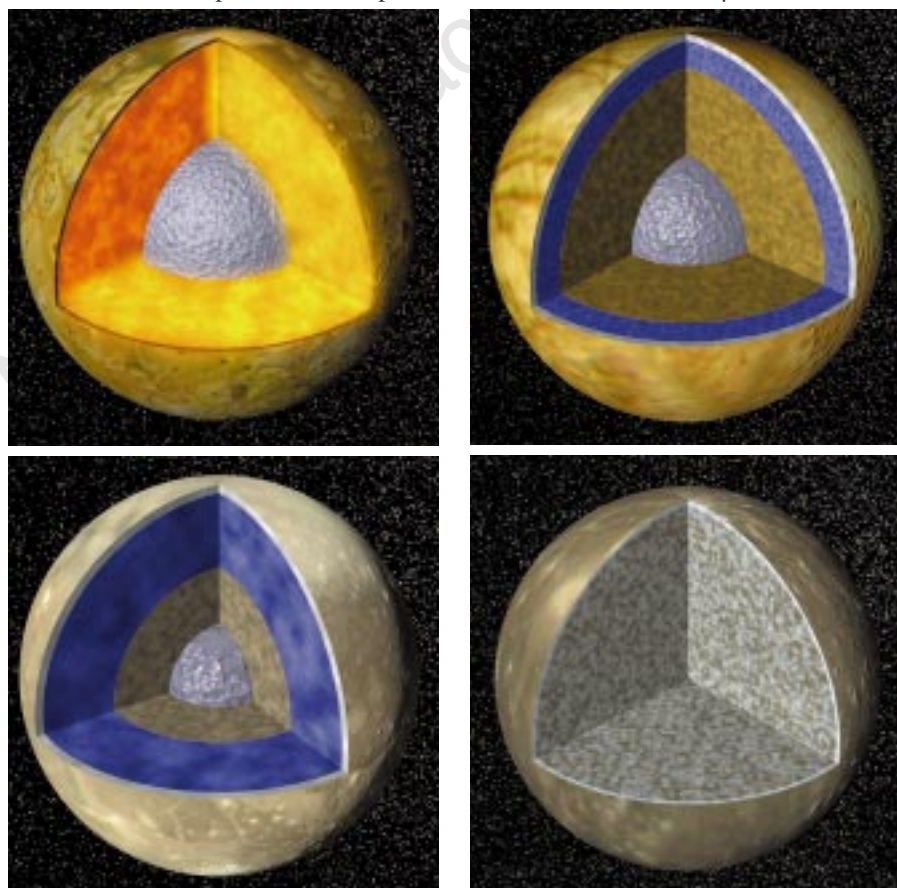


Figure 2 Cut-away drawings of the four Galilean satellites, based on gravity and magnetic data. Metal, rock and ice (where appropriate) have separated according to density in Io (top left), Europa (top right) and Ganymede. Callisto remains an undifferentiated mixture of ice and rock. (Images by Zareh Gorjian and Eric M. DeJong, JPL Digital Image Animation Laboratory.)

the inner nebula, accounting for Io's lack of ice and Europa's modest endowment³. At the positions of Ganymede and Callisto, water ice condensed, but 'supervolatile' ices such as methane and N₂ did not, because the outer protojovian nebula would have been kept warm by the surrounding solar nebula.

Tides and migration

Jupiter also influences the levels of geological activity on its satellites through the tides it raises, which flex the solid moons, and so heat them. So powerful are these tides that their heating usually overwhelms the heating due to the decay of uranium, potassium-40 and thorium in the rock of these worlds. The tides occur because the satellite orbits are eccentric (rather than circular), and the tides would tend to reduce the orbital eccentricities—but these are maintained, for the inner three satellites, by mean-motion resonances: Europa's orbital period is nearly twice that of Io, and Ganymede's period is nearly twice that of Europa. This is collectively termed the Laplace resonance. The resonance could be primordial, dating back to the time of the satellites' accretion, but mean-motion resonances are not characteristic of other satellite systems (those of Saturn, Uranus or Neptune). So the orbits of Io, Europa and Ganymede probably evolved into their present resonant configuration.

We expect that such orbital evolution is possible, because the tides that the satellites raise on Jupiter exert torques that cause the satellites' orbits to expand over geological time. Io has evolved outwards faster than Europa, and Europa faster than Ganymede. Given enough dynamical evolution, it was all but inevitable that the Laplace resonance would form (ref. 4, for example). And dynamically speaking, the Galilean satellites are an extraordinarily ancient system. If we

Table 1 Basic properties of the moons

	<i>Io</i>	<i>Europa</i>	<i>Ganymede</i>	<i>Callisto</i>
Distance from Jupiter ²⁶ (Jupiter radii)	5.9	9.4	15.0	26.4
Size ²⁷ (km)	1,821	1,565	2,635	2,405
Density ^{7,9,10,17} (g cm ⁻³)	3.53	3.02	1.94	1.85
Orbital period ²⁶ (days)	1.77	3.55	7.16	16.69
I/MR^2 (refs 7,9,10,16)	0.378 ± 0.007	0.347 ± 0.014	0.311 ± 0.003	0.406 ± 0.030

were to scale outermost Callisto to the position of Neptune, the Galilean satellites would take 16 trillion years to reach their present dynamical age! A lot can happen in such an eternity, and from this perspective the unusual body is the only one not entangled in any resonance—Callisto.

The intensity of tidal heating goes as the square of orbital eccentricity (which ranges from zero for a circle to one for a parabolic orbit) but inversely as a large power of the orbit size. So Io receives a prodigious amount of tidal heating⁵, enough to keep its volcanic surface in a more or less permanent state of upheaval⁶. Europa is less intensely heated, but tidal energy may be responsible for maintaining Europa's bright, flat and young ice surface⁵. Ganymede is farthest from Jupiter of the three resonant satellites, and as the most massive of the Galilean moons it acts as a kind of gravitational anchor for the others, so its eccentricity is not at present being resonantly forced strongly enough to provide much tidal heating. Yet numerical simulations of the satellites' orbital history suggest that during a previous resonant configuration Ganymede's eccentricity may have been nearly as high as 0.1, and the resulting tidal heating at that time could have been substantial⁴. It is tempting to link Ganymede's past tectonic and volcanic activity with this tidal heating episode⁴. Callisto, uninvolved in any of this, and with a nearly circular orbit, is negligibly heated.

On the outside looking in

What new insights has the Galileo spacecraft brought? Unheralded in comparison with the eagerly awaited high-resolution imaging and the atmospheric probe, some of the most important results derive from measurements of gravity and magnetism. Powerful constraints on the internal structure of all four Galilean satellites have been obtained.

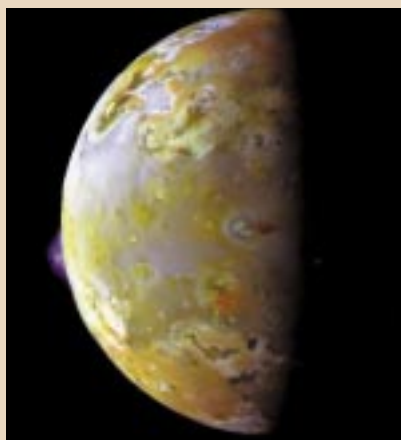
Although Newton showed that any spherically symmetric arrangement of matter behaves gravitationally as a point source, if the arrangement is distorted by rotation or tides, asymmetries are created in the gravity field. These can be measured if a spacecraft passes close enough to a body, as Galileo has, and broken down into their geometrical components—among which the quadrupole terms dominate.

Assuming that the body is in hydrostatic equilibrium⁷, the quadrupole terms can then be used to find the moment of inertia I of the body, and thus give a hint at its structure. Different radial-density profiles can result in the same moment of inertia, but as long as physically and chemically sensible configurations are chosen *a priori*⁸, the implied structure should be reasonably close to the truth. So the satellites are assumed to have layers of water ice (except for Io), silicate rock, and metal (predominantly iron), or various mixtures thereof, with the density of each layer increasing from the outside in.

Based on one close pass by Io, and two apiece by Ganymede and Europa, the moment-of-inertia factors I/MR^2 (where M and R are the satellite mass and radius, respectively) for these satellites have been determined to be 0.378 ± 0.007 , 0.311 ± 0.003 , and 0.347 ± 0.014 , respectively^{7,9,10} (Table 1). Given that I/MR^2 for a uniform sphere is 0.4, this is unambiguous evidence that all three bodies are centrally condensed. That is, they have differentiated into distinct layers of material, like the rocky planets. Io appears to have a metallic core and rock mantle⁷, which is not surprising given Io's level of tidal heating, but the result for Ganymede is particularly striking. Its I/MR^2 is the lowest yet obtained for any solid body in space, and implies that Ganymede is strongly differentiated, with its ice forming an 800–900-km mantle above a rock and metal core, with little or no room for any intervening mixed ice and rock layer. Furthermore, Ganymede has been found to possess a relatively strong dipole magnetic field, and the most physically plausible explana-

Io

Io is the most geologically and geophysically active body in the Solar System. Io is strongly tidally heated, and undergoes more-or-less-continuous volcanic resurfacing⁶. Plumes, hydrodynamically more akin to geysers, are the most spectacular volcanic phenomena (one can be seen erupting on the limb in this view from the Galileo spacecraft, and another, called Prometheus, on the line between night and day), but a wide range of volcanic landforms are found, such as vents, calderas (collapsed craters) and flows, along with mysterious, isolated mountains up to 10 km high. Heat from individual eruptions can be monitored from Earth, and Galileo has now measured both the heat and visible light from various eruptive centres; eruption temperatures can reach 1,100–1,500 K, indicating silicate volcanism⁶. The surface is covered, however, with sulphur



and sulphur compounds, including SO₂, frost. The heat flow from Io⁶ is about 2.5 mW m⁻², nearly 30 times the terrestrial average.

tion is dynamo action in a convecting, liquid portion of an inner metallic core^{11,12}.

The moments of inertia measured during the two Europa encounters differ, but even the larger value (listed in Table 1) implies that Europa, too, is quite centrally condensed¹⁰. This indicates a deep (~100–200 km) ice layer¹³, a dehydrated rocky interior, and probably an inner metallic core. Magnetic-field measurements at Europa are hard to interpret because of the strength of Jupiter's magnetic field closer to the planet, but an internally generated dipole field (implying a metallic inner core) is a possibility¹⁴.

But Callisto bucks this trend. There, all is quiet on the magnetic front¹⁵, and its I/MR^2 is consistent with an undifferentiated interior¹⁶. The lack of an intrinsic magnetic field is also consistent with, but does not require, an undifferentiated interior — as Venus and Mars demonstrate. How Callisto could have remained undifferentiated is a great puzzle (see below).

Interestingly, an undifferentiated Callisto would have an ice/dry-rock mass ratio of ~55/45, whereas a differentiated Ganymede, despite its similar size and density, has the opposite ratio of ~45/55 (this is possible because much of the ice in Callisto would be deep inside the planet, compressed into dense phases)¹⁷. This is a substantial difference in composition, not just in surface appearance, between the two bodies.

All these worlds but one

The new gravity and magnetic-field data seem to fit together in a rather neat package. All of the tidally heated satellites have separated their rock and ice, and the rock has fur-

Ganymede

Ganymede is the largest and most massive satellite in the Solar System. Its surface is predominantly water ice, and divided more-or-less equally into two terrain types, a more cratered dark terrain and a less cratered bright terrain. Both can be seen in this 54-by-90-km Galileo image. Voyager images indicated that bright terrain replaced dark terrain by the volcanic infilling of broad rifts with cleaner ice, which were then extended and fractured to form parallel sets of ridges and valleys, or grooves²². Galileo confirmed the role of extensional faulting in grooved terrain formation and the degraded nature of the dark terrain²⁸. The dark terrain seen here is irregular, knobby and cut by faults. The nearby lanes of bright terrain contain smooth swaths, supporting a volcanic origin. The parallel lineations at the upper left are grooves, and strongly resemble 'horst and graben', an extensional formation familiar on Earth. The coupled creation of bright terrain and grooves — the main geological event in the history of Ganymede — has been variously attributed to Ganymede's differentiation, the freezing of its ice mantle, and a past era of tidal heating^{4,22}.



ther differentiated into metallic inner cores and silicate outer cores. Convection in the liquid portion of the metallic inner core of Ganymede, and possibly that of Europa and Io as well, generates a dipole magnetic field. Only Callisto, having never been tidally heated, retains its primordial undifferentiated character (Fig. 2).

Beneath this beautiful architecture, however, lie deeper mysteries. It sounds simple:

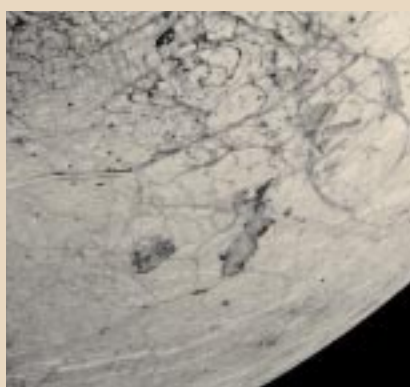
heat a rock core and the metal will melt and drain to the centre. But the rock core of an icy satellite forms when ice melts and rock falls inwards, and such a core is necessarily relatively oxidized. Under pressure and increasing core temperature, and in the presence of carbon (which must also have accreted), any iron oxides will react to join iron-bearing silicates¹⁸, leaving iron sulphide¹⁹. As a single metallic phase, its melting temperature is quite high — more than 1,500 K at core pressures. Under tidal or radiogenic heating, this means that solid-state convection in the rock sets in before melting occurs, moderating any further temperature increase and thus preventing metal-silicate separation.

Obviously, the Earth has formed its core, so it can be done. But the Earth is different in two important ways. First, it formed from materials closer to the Sun and therefore rather chemically reduced, resembling to first order the ordinary chondrite meteorites. They contain iron sulphide and free iron, and the presence of these two phases ensures that a low-melting-point (eutectic) metallic liquid can be formed. Second, the Earth was probably melted during a giant, Moon-forming impact¹⁸, enabling core formation no matter what the oxidation–reduction state of the proto-Earth. The growing Galilean satellites could also have been supplied reduced material by heliocentric (solar-orbiting) planetesimals, the small bodies from which the planets grew. Unfortunately, equilibration with ice and water during core formation should still have oxidized any rock and metal, regardless of its origin.

Furthermore, substantial accretion of

Europa

Europa's surface is almost entirely water ice. From the rarity of impact craters, its surface is young, perhaps less than 10 million years old²⁹. Relief is extraordinarily low (<200 m), which reflects both the weakness of ice as a geological material and tidal heating within the satellite. This view of a portion of the southern hemisphere (with its contrast enhanced) nevertheless shows a variety of geological units and structures. Prominent are dense assemblages of ice plates, bounded by lineaments and lanes of darker, younger ice that have accommodated the separation, rotation and translation of the plates. Long dark and bright lineaments also cross the scene, and when seen at high resolution turn out to be ridges or ridge sets²⁹. Small, dark spots may be caused by hot ice rising buoyantly from below²⁸. The irregular dark splotches at the centre of the image appear to be some type of icy volcanic flow. Evidence for an ocean on Europa, below ~10–30 km of surface ice, has been accumulating for some time²⁹. The ocean could be maintained by vigorous tidal heating in the



ice above and, to a lesser degree, in the rocky interior. Europa's ocean/ice layer is deep (~100–200 km), and the moon's interior may have differentiated into a metallic core and an overlying rock mantle^{10,14}. Thermal evolution models indicate that tidal heating may have kept the rock mantle convecting, perhaps with active volcanism at the mantle/ocean boundary¹⁹. In overall structure and activity, Europa is curiously similar to the Earth.

Callisto



Callisto is distinguished from the other Galilean satellites by showing little or no geological activity generated from within. Its ancient surface has been heavily cratered without respite, and would be monotonous were it not that the largest impacts left vast, multi-ringed fracture patterns in the crust. This Galileo image shows a close-up, 38 km across, of one of the scarps in the outer portion of the ~4,000-km-diameter Valhalla multi-ringed structure. A black cliff-shadow stretches across a ruined

landscape of icy hillocks and small craters, partially buried by dark, smooth material. This 'blanketing' is not seen on Ganymede's dark terrain. It may be the result of the loss of ice over geological time to sublimation and sputtering by magnetospheric charged particles³⁰, allowing rocky material originally bound in the ice to accumulate. This makes sense if the rock component is substantial, such as would be true for a primordial ice and rock mixture of 20–30% rock by volume¹⁷.

heliocentric planetesimals would compound another deep mystery — how to keep Callisto from differentiating. Both Ganymede and Callisto are large enough bodies that the gravitational potential energy of their formation is more than enough to have melted all their ice²⁰. Only by accreting gradually enough, or from bodies small and/or slow enough that the heat released by the collisions is not deeply buried, can accretional heat be efficiently radiated or convected (via nebular gas) away from a growing satellite surface^{3,20}. Unfortunately, heliocentric planetesimals move on high-velocity, hyperbolic trajectories with respect to the satellites, and release one to two orders of magnitude more energy upon collision than local planetesimals orbiting Jupiter — a large enough heliocentric impact could even trigger wholesale (satellite-wide) ice–rock differentiation²¹. No one has yet come up with a convincing mechanism to prevent ice melting and differentiation within Callisto^{20,22}.

The implications of an undifferentiated Callisto reach further. As discussed above, the conventional explanation of the satellites' compositional gradient involves the luminosity of proto-Jupiter, but this assumes that the satellites accreted not far from their present orbital distances. This may not be true in general for satellites or for planets. Evidence is accumulating that the planets migrated great distances inward or outward as they formed²³, and inward migration may explain the 'hot Jupiters' around other stars, discovered over the past few years²⁴. And because of the high density of the protojov-

ian nebula, gas drag or other dynamic satellite–nebula interactions could have caused the satellites to spiral in towards Jupiter after they formed. The satellites we see today would have then been left stranded when the nebula dissipated, and may only be the last in a long line of worlds, most of which were consumed by proto-Jupiter^{3,20}.

In this picture, each satellite presumably accreted from a similar mix of ice and rock as it migrated, so Io and Europa would have needed some way to shed unwanted ice. The late E. M. Shoemaker hypothesized that the early bombardment of Jupiter's system (more than four billion years ago) would have been fierce enough to completely strip Io of an original ice mantle and remove most of Europa's as well. But an undifferentiated Callisto imposes stringent conditions on the nature of any such bombardment, leaving the whole 'orbital migration' picture in question. Yet from a theoretical standpoint it is difficult to make the young Galilean satellites sit still in their nebular nursery³. Perhaps there were hardly any comets large enough to melt Callisto. Or perhaps Callisto is differentiated after all, but is fooling us with non-hydrostatic components in its gravity field¹⁷. There is as yet no attractive solution to the mystery of Callisto.

A mission from the Pope

Last winter a meeting called "The Three Galileos" was held in Padua, Italy, to celebrate the man, the spacecraft and the new Italian national telescope. A highlight was the official presentation of a volume of

images from the Galileo Project to Pope John Paul II. The Pope was reported to be interested and perceptive, and encouraged continued exploration of the Universe²⁵. Happily, Galileo (the spacecraft) should in this instance have no trouble acceding to papal wishes. The Orbiter completes its nominal mission this December, but the US space agency NASA plans to let it continue, taking advantage of the final orbital geometry to repeatedly encounter Europa, and then, by using Callisto for several gravity assists, to encounter Io. Much more of Europa will be imaged at high resolution, and Io will be for the first time. Moreover, the additional gravity and magnetic field measurements should greatly improve the interior models, providing a richer portrait of the satellites' geophysical personalities. Understanding the origin and evolution of the Galilean satellites will doubtless take time and require further explorations, but it will be a vital element in building a comprehensive theory of planet formation. □

William B. McKinnon is in the Department of Earth and Planetary Sciences and McDonnell Center for the Space Sciences, Washington University, St Louis, Missouri 63130, USA.

- Galilei, G. *The Starry Messenger* (1610) in *Telescopes, Tides, and Tactics* (ed. & transl. Drake, S.) 89 (Univ. Chicago Press, 1983).
- Shoemaker, E. M. & Wolfe, R. F. in *Satellites of Jupiter* (ed. Morrison, D.) 277–339 (Univ. Arizona Press, Tucson, 1982).
- Lunine, J. I. & Stevenson, D. J. *Icarus* **52**, 14–39 (1982).
- Showman, A. P. & Malhotra, R. *Icarus* **127**, 93–111 (1997).
- Cassen, P. M., Peale, S. J. & Reynolds, R. T. in *Satellites of Jupiter* (ed. Morrison, D.) 93–128 (Univ. Arizona Press, Tucson, 1982).
- Spencer, J. R. & Schneider, N. M. *Annu. Rev. Earth Planet. Sci.* **24**, 125–190 (1996).
- Anderson, J. D., Sjogren, W. L. & Schubert, G. *Science* **272**, 709–712 (1996).
- Freud, S. *New Introductory Lectures in Psychoanalysis* (1933) (ed. & transl. Strachey, J.) 32 (W. W. Norton, New York, 1965).
- Anderson, J. D., Lau, E. L., Sjogren, W. L., Schubert, G. & Moore, W. B. *Nature* **384**, 541–543 (1996).
- Anderson, J. D., Lau, E. L., Sjogren, W. L., Schubert, G. & Moore, W. B. *Science* **276**, 1236–1239 (1997).
- Stevenson, D. J. *Nature* **384**, 511–512 (1996).
- Schubert, G., Zhang, K., Kivelson, M. G. & Anderson, J. D. *Nature* **384**, 544–545 (1996).
- Ransford, G. A., Finnerty, A. A. & Collerson, K. D. *Nature* **289**, 21–24 (1981).
- Kivelson, M. et al. *Science* **276**, 1239–1241 (1997).
- Khurana, K. K. et al. *Nature* **387**, 262–264 (1997).
- Anderson, J. D., Lau, E. L., Sjogren, W. L., Schubert, G. & Moore, W. B. *Nature* **387**, 264–266 (1997).
- McKinnon, W. B. *Icarus* (in the press).
- Arculus, R. J. et al. in *Origin of the Earth* (eds Newsom, H. E. & Jones, J. H.) 251–271 (Oxford Univ. Press, New York, 1990).
- McKinnon, W. B. *EOS Trans AGU* **77**, Fall Meet. Suppl., F442 (1996).
- Stevenson, D. J., Harris, A. W. & Lunine, J. I. in *Satellites* (eds Burns, J. A. & Matthews, M. S.) 39–88 (Univ. Arizona Press, Tucson, 1986).
- Tonks, W. B., Pierazzo, E. & Melosh, H. J. *Icarus* (in the press).
- McKinnon, W. B. & Parmentier, E. M. in *Satellites* (eds Burns, J. A. & Matthews, M. S.) 716–763 (Univ. Arizona Press, Tucson, 1986).
- Malhotra, R. *Astron. J.* **110**, 420–429 (1995).
- Lin, D. N. C., Bodenheimer, P. & Richardson, D. C. *Nature* **380**, 606–607 (1996).
- www.jpl.nasa.gov/galileo/status970212.html
- Burns, J. A. in *Satellites* (eds Burns, J. A. & Matthews, M. S.) 1–38 (Univ. Arizona Press, Tucson, 1986).
- Davies, M. E. et al. *Celest. Mech. Dynam. Astron.* **53**, 377–397 (1991).
- Belton, M. J. S. et al. *Science* **274**, 377–391 (1996).
- Carr, M. H. et al. *Nature* (in the press).
- Moore, J. M., Mellon, M. T. & Zent, A. P. *Icarus* **122**, 63–78 (1996).

Derges's designs

Drop sheets of photographic paper into a river and let the play of light and water paint its own picture. That is one of the innovative techniques used by Susan Derges in her 'photo-works'.

Martin Kemp

Nature is an alarmingly tough competitor for any artist. It parades an endless succession of beguiling and elusive effects whose subtlety and vitality present an enduring challenge to the inert media available to the painter and sculptor.

In particular, the capturing of the transitory beauties of light and motion in water has presented perpetual problems, to be faced by the Renaissance painter of the *Baptism of Christ* or *Birth of Venus* no less than by an Impressionist seeking to evoke a seascape saturated with sunlight and fractured by scattered reflections of rocks and trees.

One solution to the problem of entrapping such natural effects is brilliantly realized by the Devon artist, Susan Derges. She collaborates directly with the natural phenomena themselves, in such a way that nature inscribes its own order and disorder on the designated surfaces of the works of art.

Derges is internationally admired for her photo-works in which nature imprints patterns and rhythms of motion, growth and form directly on the light-sensitive surface of the photographic emulsion. She has collaborated intensively with such phenomena as falling water drops, busy honeycombs and vessels of germinating toad's eggs. Such images make us look again — look really hard — in keeping with a historical succession that runs through Leonardo, Dürer, Goethe and Ruskin. Her most recent work tackles the motion of water, that hell of the predictive sciences.

In one series she is treating the surface of a river, the Taw in Devon, as a photographic transparency, searching out flow patterns that remain relatively stable. Sheets of photographic paper in a large slide (about 6 feet by 2 feet) are immersed just below the surface at night and illuminated by a brilliant but diffuse flash lasting a microsecond. The colour images register the peaks of the wavelets as white lines and the troughs as tinted by moonlight and reflected vegetation. The black-and-white photographs, such as the one illustrated here, are made with a negative process which inverts the tonality. The crests accordingly are represented as a rippling cellular network of dark membranous filaments.

In a technical sense, a student of hydraulics might scrutinize her large prints as lucid illustrations of wave trains, capillary waves and so on. But they are more than technical illustrations. They suggestively



Susan Derges's *Untitled Photogram (Riusr Tor, 1966)* (inset) and detail (main picture).

evoke the pattern-forming propensities of physically diverse phenomena, not only those exhibited dynamically by fluids but also those discernible in the growth of cells, the wrinkling of tissues and so on — across a wide range of scales.

At one level, the images are life-size representations of very specific events. And they are

unusually large for photographs. Yet, at another level, these and others of her works aspire to reveal scale-less commonalities in nature, both as sensed intuitively and as investigated mathematically within chaos theory. □

Martin Kemp is in the Department of the History of Art, University of Oxford, 35 Beaumont Street, Oxford OX1 2PG, UK.

Of fingers, toes and penises

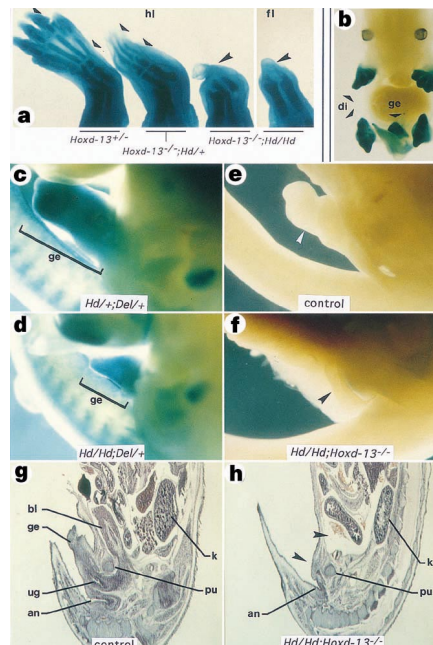
Vertebrate *Hox* genes are essential for limb development. The posterior-most *Hoxd* and *Hoxa* genes are required for growth and patterning of digits and are also strongly expressed in the genital bud, which gives rise to the urogenital system, including the penis. Here, we show that removal of posterior *Hox* gene function results in a concomitant loss of digits and genital bud-derivatives, illustrating that similar developmental mechanisms are at work in these different buds.

We raised compound mutant mice carrying the *hypodactyly* (*Hd*) allele¹ and null alleles of the *HoxD* complex inactivating either *Hoxd-13* (ref. 2), or *Hoxd-11*, *Hoxd-12* and *Hoxd-13* in *cis* configuration *HoxD^{Del}* (ref. 3). *Hd* mice have a deletion in *Hoxa-13* (ref. 1), so these combinations removed most *Hox* gene function in the digits. The effect was comparable to that observed when using a *Hoxa-13* loss-of-function allele⁴: a total digit agenesis (Fig. 1a). Double-mutant limbs were truncated right after the zeugopods (Fig. 1a; large arrowheads) with only traces of proximal tarsal condensations (Fig. 1a).

Expression of *Hoxd* genes was visualized using a *lacZ* reporter gene inserted either in *Hoxd-13* (Fig. 1b) or at the *HoxD^{Del}* locus³. β -gal staining revealed a sizeable blue eminence in control *Hd/+; Del/+* mice (Fig. 1c) which showed considerable digit development. But a second *Hd* allele (*Hd/Hd; Del/+* mice) induced a severe truncation of the eminence (Fig. 1d). Removing group-13 function (*Hd/Hd; Hoxd-13^{-/-}*) led to complete agenesis of the bud (Fig. 1e, f). This progressive reduction in genital bud size with reductions in *Hox* gene dose was accompanied by proportional reductions in digit size (compare Fig. 1a with c, d). Mild anatomical defects in these two structures also occur in the human hand-foot-genital syndrome which involves a mutation within *Hoxa-13* (ref. 5).

Histological analyses of *Hd/Hd; Hoxd-13^{-/-}* fetuses confirmed the complete absence of external genitalia (Fig. 1g, h). In addition, the lower urinary system, particularly the bladder and urethra, were missing, as well as derivatives of the urogenital sinus. Kidneys, which do not express *Hoxd-13* (ref. 6), were present in compound mutant animals. *Hd/Hd; Hoxd-13^{-/-}* fetuses died *in utero* before embryonic day 16 like most *Hd/Hd* mice¹.

The development of the distal limbs and genital eminence share important features. Both structures are regions of apical growth and involve epithelial-mesenchymal interactions. Further, digits and genitalia represent 'morphogenetic ends' of the body, digits



at the distal end of the limbs and genitalia at the trunk posterior. Molecular data support this analogy, as most genes expressed during digit development are also functional in the genital eminence. *Hox* genes are particularly relevant as 5'-located *Hoxa* and *Hoxd* genes are involved in organizing the terminal parts of axial structures⁵⁻⁷.

There is other evidence to suggest a phylogenetic link between digits and external genitalia. In fish, which lack digits and external genitalia, posterior *Hox* genes are expressed in areas from which these structures emerged (the cloacal region and paired fins). Also, *Hoxd* gene expression in developing murine digits and genital eminence is controlled by the same transcriptional enhancer element⁸⁻¹⁰. Finally, *Hoxa-13* and *Hoxb-13*, the other group-13 paralogues, are expressed neither in digits nor in the genital bud^{11,12}.

During evolution, the acquisition of the distal *Hoxd* expression domain in paired fin buds may have been linked to the emergence of digits by extending the terminal phase of proliferation¹³. Related phylogenetic histories between digits and the penis/clitoris thus seem plausible and may reflect combined adaptive or preadaptive solutions to the transition towards a terrestrial environment. Palaeontological observations do not exclude the possibility that ancestral tetrapods had protruding external genitalia. As such animals were aquatic¹⁴, a genital organ may not have been linked to terrestrial life but instead might have reflected the requirement for elaborate internal fertilization.

Digits and an ancestral genital organ could have emerged from the same genetic

Figure 1 Absence of digits and genitalia in mice lacking posterior *Hox* function. **a**, Dissected hindlimbs (hl) and forelimbs (fl) from control mice and mice lacking most or all group-13 function. Small arrowheads, digits; large arrowheads, absence of digit. **b**, Expression of *Hoxd-13/lacZ* reporter gene in digits (di) and genital eminence (ge) of a 12.5-day-old fetus (E12.5). **c**, Expression of *Hoxd/lacZ* reporter in E13.5 developing genital eminence of fetuses with a phenotype close to normal (*Hd/+; Del/+*). **d**, Removal of one dose of *Hoxa-13* (*Hd/Hd; Del/+*) led to truncation of the genital eminence. Left hindlimbs were removed to view the eminence. Other blue domains are sites of posterior *HoxD* locus expression. **e**, Control E13.5 fetus with genital eminence (arrow). **f**, Same-aged fetus in absence of group-13 function. **g,h**, Mid- to parasagittal sections through fetuses shown in **e** and **f**, respectively. ug, urogenital sinus; an, anus; k, kidneys; pu, pubis; bl, bladder. In the absence of group-13 function, the eminence is absent as well as the bladder (arrows). The rectum and anus are also abnormal (arrow).

regulatory innovation. It is nevertheless equally possible that the underlying molecular mechanisms were first deployed in forming an external genital organ and subsequently used to produce digits, or vice versa. In any event, the combined improvement of apical growth mechanisms along these axes produced the adaptive potentials necessary for efficient locomotion and internal fertilization in the terrestrial environment.

**Takashi Kondo, József Zákány
Jeffrey W. Innis, Denis Duboule**

Department of Zoology and Animal Biology,
University of Geneva, Quai Ernest Ansermet 30,
1211 Geneva 4, Switzerland
and Departments of Human Genetics and
Pediatrics,
University of Michigan Medical School,
Ann Arbor, Michigan 48109-0618, USA
e-mail: duboule@sc2a.unige.ch

- Mortlock, D. P., Post, L. C. & Innis, J. W. *Nature Genet.* **13**, 284–289 (1996).
- Dollé, P. et al. *Cell* **75**, 431–441 (1993).
- Zákány, J. & Duboule, D. *Nature* **384**, 69–71 (1996).
- Fromental-Ramain, C. et al. *Development* **122**, 2997–3011 (1996).
- Mortlock, D. P. & Innis, J. W. *Nature Genet.* **15**, 179–180 (1997).
- Dollé, P., Izpisua-Belmonte, J.-C., Brown, J. M., Tickle, C. & Duboule, D. *Genes Dev.* **5**, 1767–1776 (1991).
- Kondo, T., Dollé, P., Zákány, J. & Duboule, D. *Development* **122**, 2651–2659 (1996).
- van der Hoeven, F., Zákány, J. & Duboule, D. *Cell* **85**, 1025–1035 (1996).
- Hérault, Y., Fraudeau, N., Zákány, J. & Duboule, D. *Development* **124**, 3493–3500 (1997).
- Peichel, C. L., Prabhakaran, B. & Vogt, T. F. *Development* **124**, 3481–3492 (1997).
- Peterson, R. L., Papenbrock, T., Davda, M. M. & Awgulewitsch, A. *Mech. Dev.* **47**, 253–260 (1994).
- Zeltser, L., Desplan, C. & Heintz, N. *Development* **122**, 2475–2484 (1996).
- Sordino, P., van der Hoeven, F. & Duboule, D. *Nature* **375**, 678–681 (1995).
- Coates, M. I. *Development* (suppl.) 169–180 (1994).

Chaperoning extended life

The capacity to moderate internal and external stress is arguably the central function regulating senescence in whole-animal ageing¹⁻³. During ageing, molecular chaperones such as heat-shock proteins are thought to combat stress-related senescent dysfunction^{4,5}. In transgenic *Drosophila melanogaster*, with varying copy numbers of the gene *hsp70* encoding heat-shock protein hsp70, we found that heat-induced expression of hsp70 increased lifespan at normal temperatures. Only a brief, low level of expression was required to obtain a long-term improvement in survival.

Thermally conditioned *D. melanogaster* and *Caenorhabditis elegans* exhibit greater longevity when mildly heated adults are returned to normal temperatures^{6,7}. In *D. melanogaster*, such extended longevity is caused by an immediate increase in age-specific survival that persists for several weeks⁶. This change in survival is not easily explained by the removal of mortality costs of reproduction because these levels of thermal conditioning do not reduce egg laying. Rather, improved longevity may represent the modulation of age-related stress by the induced expression of heat-shock proteins.

We used transgenic *D. melanogaster* strains (from S. Lindquist), which varied in the number of copies of the inducible *hsp70* gene^{8,9}, to determine the effect of hsp70 protein on survival during ageing at normal temperatures. These allelic strains control

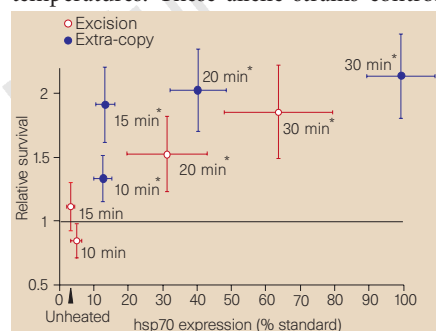


Figure 1 Effect of hsp70 expression on relative survival in 'extra-copy' and 'excision' strains of *Drosophila*. Within each strain, relative survival is calculated over the 14 days after heat shock compared with a control cohort of unheated flies (see ref. 6). Life tables were constructed from cohorts of about 1,200 flies per treatment. Heat treatments varied from 10 to 30 min, and durations that produced significant increases in survival are marked with asterisks. Levels of hsp70 protein were measured by enzyme-linked immunosorbent assay (ELISA)⁹. Arrow shows the background hsp70 level in unheated flies. Standard errors are shown for estimates of both variables.

for the effects of insertional mutagenesis. The 'excision' strain carried the wild-type complement of constitutive *hsp70* and heat-inducible *hsp70* genes. The 'extra-copy' strain added to the second chromosome a total of 12 additional copies of heat-inducible *hsp70*. The excision strain carried only a remnant P-element construct at the same integration site.

Four days after emergence of adult males, we heat-treated each strain with a brief, non-lethal 36 °C pulse of varying duration, and measured the subsequent expression of induced hsp70 protein, age-specific survival and remaining life expectancy at 24 °C. In independent trials we replicated these results, characterized the time-course of hsp70 protein induction, and observed the effects of longer heat treatments.

The presence of hsp70 protein increased subsequent survival at normal temperatures when expression exceeded 10–12% of standard (Fig. 1). The improvement in survival reached a plateau with higher levels of expressed protein. 'Extra-copy' but not 'excision' flies heated for 10 or 15 min expressed hsp70 and had improved survival over the two-week period after heat shock. Under these conditions, life expectancy increased by as much as 7.9% in 'extra-copy' flies, which is a substantial rise given the prevailing low mortality rates of young adults. Short heat treatments failed to induce hsp70 or to improve survival in the excision flies. These observations show that hsp70 moderates survival during subsequent ageing.

Heat treatments of 20 and 30 min induced hsp70 to levels greater than 30% of standard in both strains and improved survival by similar amounts in each. Because the relationship between improved survival and hsp70 is similar in 'extra-copy' and 'excision' strains, we infer that the strains differ solely in their ability to express hsp70 protein for a given heat dose. This functional similarity also suggests that *hsp70* copy number, and not merely differences in insert size, causes the improved survival after heat shock.

The whole-fly titre of induced hsp70 protein is transient, but its effect on age-specific survival is persistent. A transient expression of a molecular chaperone may increase age-specific survival through its ability to renature, assemble and disassemble many non-heat-shock proteins¹⁰, and to interact with other stress-response mechanisms such as superoxide dismutase⁵. It is also possible that hsp70 persists at high levels in some critical and specific tissue over the period of increased survival. Although hsp70 and heat-shock proteins in general are tightly regulated¹⁰, transient but effective levels of hsp70 could be present when stress is routinely encountered, and hsp70

may repair and restore higher order cell functions which themselves would otherwise accelerate senescence.

Marc Tatar*, Aziz A. Khazaeli

James W. Curtsinger

Department of Ecology, Evolution and Behavior,
University of Minnesota, St Paul,
Minnesota 55108, USA

e-mail: Marc_Tatar@Brown.edu

*Present address: Department of Ecology and Evolutionary
Biology, Box G-W, Brown University, Providence, Rhode Island
02912, USA.

1. Jazwinski, S. M. *Science* **273**, 54–59 (1996).
2. Lithgow, G. J. *BioEssays* **18**, 809–815 (1996).
3. Murakami, S. & Johnson, T. E. *Genetics* **143**, 1207–1218 (1996).
4. Heydari, A. R., Takahashi, T., Gütsmann, A., You, S. & Richardson, A. *Experientia* **50**, 1092–1098 (1994).
5. Wheeler, J. C., Bieschke, E. T. & Tower, J. *Proc. Natl Acad. Sci. USA* **92**, 10408–10412 (1995).
6. Khazaeli, A. A., Tatar, M., Pletcher, S. D. & Curtsinger, J. W. *J. Gerontol. A* **52**, B48–B52 (1997).
7. Lithgow, G. J., White, T. M., Melow, S. & Johnson, T. E. *Proc. Natl Acad. Sci. USA* **92**, 7540–7544 (1995).
8. Welte, M. A., Tetraault, J. M., Dellavalle, R. P. & Lindquist, S. L. *Curr. Biol.* **13**, 842–853 (1993).
9. Feder, M. E., Cartano, N. V., Milos, L., Krebs, R. A. & Lindquist, S. L. *J. Exp. Biol.* **119**, 1837–1844 (1997).
10. Parsell, D. A. & Lindquist, S. *Annu. Rev. Genet.* **27**, 437–496 (1993).

Probabilities of conspecificity

G. Suwa *et al.*¹ and E. Delson's News and Views² address one of the most contentious issues in palaeoanthropology, that of the boundaries (if any) between hominid species, in the context of an exciting new fossil from Konso in Ethiopia. Suwa *et al.* attribute this specimen to *Australopithecus boisei*, but note that it has some similarities to the South African robust australopithecines first described by Robert Broom more than 50 years ago.

As the *Nature* authors say, there is certainly a need to be careful about 'pigeon-holing' particular fossils into a particular species. One way of addressing this issue is to try to assess probabilities of conspecificity when pairs of specimens are compared. By using a simple morphometric technique based on comparison of anatomical measurements of conspecific pairs of extant taxa (1,260 specimens, representing 70 vertebrate and invertebrate species), we have found a lognormal distribution of the standard error of the *m*-coefficient associated with equations of the general form $y = mx + c$ derived from linear regression analysis³. The method has been used in comparisons between the type specimens of two 'robust' australopithecines, OH 5 (*A. boisei* from Olduvai Gorge in Tanzania, described by P. V. Tobias), and TM 1517 (*A. robustus*, from Kromdraai in South Africa), suggesting a high probability of conspecificity. The log standard error value of -1.31 determined by comparing cranial

dimensions measurable from both specimens is within two standard deviations of the mean value of -1.78 obtained from conspecific comparisons of large samples of extant vertebrate and invertebrate taxa³.

Although this result suggests that the type specimens of *A. robustus* and *A. boisei* have a high probability of conspecificity, this does not mean that all specimens attributed to *A. boisei* should necessarily be referred to as *A. robustus*. As a South African in the new South Africa, I believe we should move away from rigid application of the binomial system of classification in palaeontological contexts.

J. F. Thackeray

Transvaal Museum, PO Box 413,
Pretoria 0001, South Africa

1. Suwa, G. et al. *Nature* **389**, 489–492 (1997).

2. Delson, E. *Nature* **389**, 445–446 (1997).

3. Thackeray, J. F. et al. *S. Afr. J. Sci.* **93**, 195–196 (1997).

Xenoturbella's molluscan relatives...

Despite detailed morphological studies^{1–4}, the phylogenetic relationships of *Xenoturbella bocki* Westblad 1949 have remained unclear. The marine, worm-like *X. bocki* was first described as an acoel flatworm⁵. Later it was proposed to be a deuterostome¹, and most recently as the sister taxon of the Bilateria⁶. Here we present DNA sequence data that place *X. bocki* within the protostome clade Eutrochozoa.

We used standard DNA extraction, polymerase chain reaction (PCR) and sequencing techniques to sequence 1,759 nucleotides of the small-subunit ribosomal RNA gene (18S rRNA) and 708 base pairs of the protein-coding mitochondrial cytochrome *c* oxidase subunit I gene (COI) from five specimens of *X. bocki* collected on the west coast of Sweden. We sequenced the corresponding COI fragment from the flatworm *Graffilla buccinicola* for comparison. We used these and sequences obtained from GenBank to construct two matrices for cladistic analysis.

The 18S rRNA matrix (778 informative characters) comprised 47 sequences from 18 phyla relevant to the position of *X. bocki* of which 35 sequences were previously aligned according to secondary structure⁷. The remaining sequences were added using ClustalW⁸. There was less COI data — the COI matrix (370 informative characters) comprised 22 taxa from 8 phyla.

We used parsimony jack-knifing (resampling with deletion of a fraction of the characters) to generate phylogenetic hypotheses and to evaluate their support (using pre-release 'Xac' software, J. S. Farris)⁹. To eliminate homoplasious nucleotide positions in

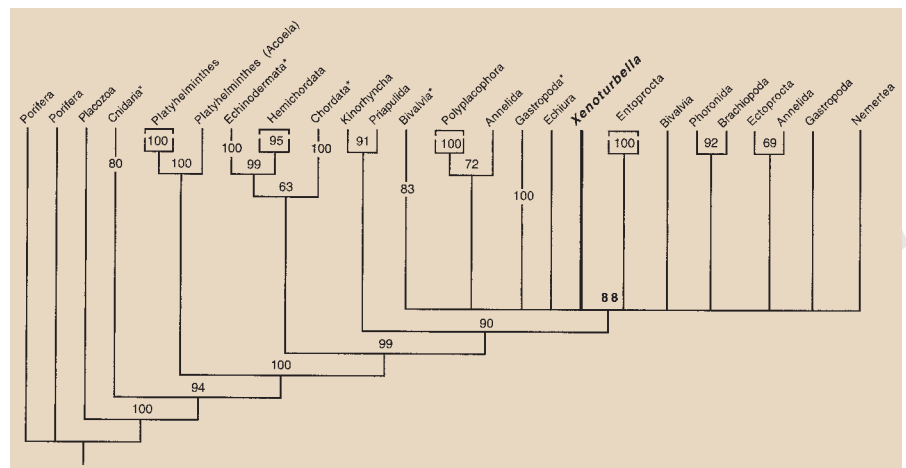


Figure 1 Consensus tree showing groups present in 60% of jack-knife replicates from analysis of 18S rRNA matrix (successive weighting of characters, 10 iterations with 100 replicates each, deletion frequency e^{-1}). Labels indicate jack-knife frequencies. Clades marked with asterisks represent multiple terminals.

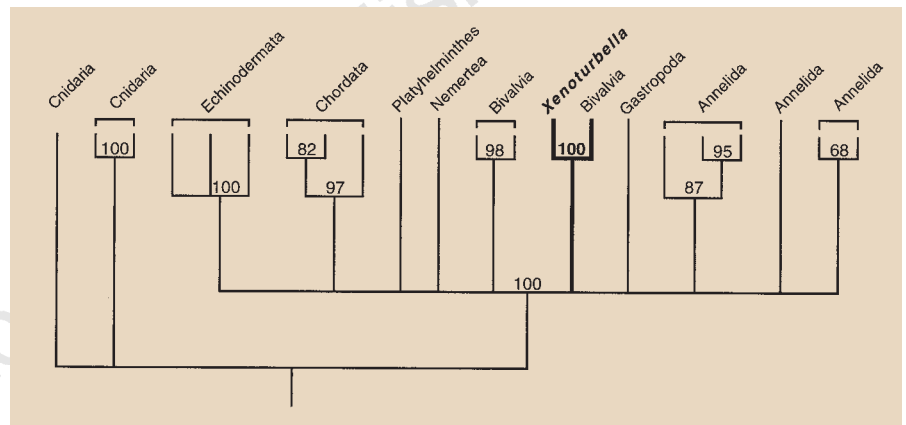


Figure 2 Consensus tree showing groups present in 60% of jack-knife replicates from analysis of COI matrix (3,000 replicates, 5 random additions and branch swapping, deletion frequency e^{-1}). Labels indicate jack-knife frequencies. Full details of tree topology and sequence alignment are available from the authors.

the 18S rRNA sequences, we combined parsimony jack-knifing with successive weighting of characters¹⁰. We analysed the COI matrix without successive weighting.

The majority-rule consensus tree (Fig. 1) shows groups present in more than 60% of the jack-knife replicates in the analysis of 18S rRNA data. *X. bocki* is part of a highly supported (88%) unresolved clade consisting of representatives of Mollusca, Annelida (including Pogonophora), Echiura, Phoronida, Brachiopoda, Entoprocta, Ectoprocta and Nemertea. The hemichordates, previously suggested as close relations of *X. bocki*¹ are the sister group of echinoderms (99%) within a deuterostome clade (63%). Monophyly of the flatworms, including the acoel *Praesagittifera*, is supported in all jack-knife replicates, and their sister-group relationship to other bilaterian taxa is also highly supported (94%). These results are not compatible with hemichordates or platyhelminths as sister groups of *X. bocki*, but suggest that *X. bocki* is a member of the Eutrochozoa (non-moulting protostomes), a clade proposed in previous studies of metazoan evolution on the basis of 18S rRNA data^{11,12}.

The 60% majority-rule consensus tree resulting from jack-knife analysis of the COI data (Fig. 2), places *X. bocki* as part of the Bilateria forming a clade with the protobranch bivalve mollusc *Ennucula tenuis* in all 3,000 jack-knife replicates. The *X. bocki* COI sequence was translated into amino acids using universal, vertebrate, generalized invertebrate (mollusc), echinoderm, ascidian and flatworm mitochondrial translation codes. However, analysis of the amino-acid sequences did not improve tree resolution, and we found no unambiguous stop codons with any of the codes.

The COI gene appears to evolve too rapidly to be useful in analyses of relationships at phylum or higher levels. Nonetheless, there is a strong indication of a close relationship between *X. bocki* and the bivalve *E. tenuis*, a clade compatible with the 18S rRNA results and supported by similarities in the oogenesis of protobranch bivalves and *X. bocki* described by Israels¹³. The same clade (79%) was present as part of a protostome clade (84%) excluding flatworms, after analysis without successive weighting of a combined 18S rRNA and

COI matrix (3,000 replicates with branch swapping, 5 random additions). Thus both data sets analysed independently and in combination support *Xenoturbella bocki* as a member of the Eutrochozoa.

Michael Norén, Ulf Jondelius

Swedish Museum of Natural History,
POB 50007, SE-104 05 Stockholm, Sweden
e-mail: ulfj@nrm.se

1. Reisinger, E. Z. *Wiss. Zool.* **164**, 188–198 (1960).
2. Franzén, A. & Afzelius, B. A. *Zool. Scr.* **16**, 9–17 (1987).
3. Pedersen, K. J. & Pedersen, L. R. *Acta Zool.* **67**, 103–113 (1986).
4. Pedersen, K. J. & Pedersen, L. R. *Acta Zool.* **69**, 231–246 (1988).
5. Westblad, E. *Ark. Zool.* **1**, 11–29 (1949).
6. Ehlers, U. & Sopott-Ehlers, B. *Zoomorphology* **117**, 71–80 (1997).
7. van de Peer, Y., Jansen, J., de Rijk, P. & de Wachter, R. *Nucleic Acids Res.* **25**, 111–116 (1997).
8. Thompson, J. D., Higgins, D. G. & Gibson, T. J. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
9. Farris, J. S., Albert, V. A., Källersjö, M., Lipscomb, D. & Kluge, A. G. *Cladistics* **12**, 99–124 (1996).
10. Farris, J. S. *Syst. Zool.* **18**, 374–375 (1969).
11. Winneppenninckx, B., Bäckeljau, T. & De Wachter, R. *Mol. Biol. Evol.* **12**, 641–649 (1995).
12. Winneppenninckx, B., Bäckeljau, T. & De Wachter, R. *Mol. Biol. Evol.* **13**, 1306–1317 (1996).
13. Israelsson, O. *Nature* **390**, 32 (1997).

...and molluscan embryogenesis

Xenoturbella bocki Westblad¹ is a strange animal — a 2-cm-long, slowly moving ciliated bag with no anus and no organs except for a position-sensing statocyst containing flagellated statoconia². Despite the animal's peculiarities, it has been neglected by most textbooks. I now report a study of oogenesis in *X. bocki* which, together with the nucleotide data of Norén and Jondelius³, contradicts earlier hypotheses as to the phylogeny of the animal and instead suggests a molluscan relationship close to or within the protobranch bivalves.

Because of its simplicity in organization, *X. bocki* has been interpreted as one of the most primitive metazoa^{4,5}, a neotenous deuterostome⁶, or a relation of acoelomorph flatworms^{1,7}, although the latter two hypotheses have recently been rejected^{8,9}. Its proposed position as an early offshoot from the metazoans has brought *X. bocki* into the spotlight, as this might significantly influ-

ence our understanding of metazoan radiation and evolution. Norén and Jondelius's study³ of 18S ribosomal RNA and cytochrome oxidase I (COI) nucleotide sequences rejects all earlier hypotheses and suggests a molluscan relationship, perhaps a relationship with protobranch bivalves. Morphological data neither support nor reject this hypothesis^{1,5,7,9–11}.

The oogenesis of *X. bocki* has been described only briefly¹. The oogonia have nuclei with a peripheral rim of heterochromatin and a single, homogeneous nucleolus. The chromatin becomes dispersed, and a previtellogenic oocyte is formed. Most oocytes continue to grow in the parenchyma, whereas others develop within the gastrodermis close to the parenchyma. The nucleus has amoeboid processes. After a growth period, the nucleus becomes rounded again and vacuoles are formed in the centre of the nucleolus. The nucleoli seem to migrate to the periphery before vitellogenesis.

The oocyte attaches to the gastrodermis, yolk granules begin to accumulate, and the nucleolus becomes homogeneously basophilic and disintegrates into numerous micronucleoli so that the main nucleolus completely disappears. The micronucleoli arrange around the periphery beneath the nuclear envelope opposite to the attachment area of the ovum (Fig. 1a), persisting until the ovum becomes mature. The mature ova, which were not found by Westblad¹, have irregularly rounded nuclei without any remaining nucleoli. There are no nurse or nutritive cells.

The arrangement of nucleolar vacuoles is similar to that of most molluscs and sipunculoids, and micronucleoli are known from different metazoans but not from placozoans, poriferans, cnidarians, or acoelomorph flatworms (refs 12, 13 and references therein; data not shown). However, they are dissolved before the end of vitellogenesis or, if they persist, they remain scattered within the nucleus or along the whole nuclear envelope. In protobranch bivalves on the other hand, in which oogenesis has not yet been described, as well as a main nucleolus, micronucleoli are present and are arranged along one end of the germinal

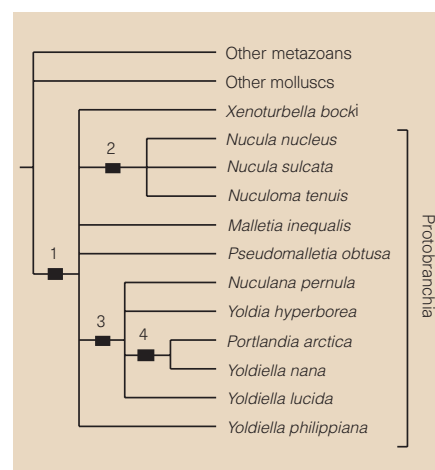


Figure 2 Cladistic analyses of oogenesis indicates that *Xenoturbella bocki* is a sister group or a subgroup of protobranch bivalves. The analysed characters, with their apomorphic states, are: **1**, micronucleoli migrate at one end of the germinal vesicle; **2**, few, large micronucleoli; **3**, chromatin is not scattered but accumulated at the same end as the micronucleoli; and **4**, large, red lipid droplets are present in the ooplasm. The molluscan data include information from refs 12, 13, and references therein, and from observations of 2 caudofoveatans, 3 solenogastres, 2 polyplacophores, 29 gastropods, 3 scaphopods, and 24 bivalves.

vesicle. They also persist throughout vitellogenesis (Fig. 1b). This character is a synapomorphy that is shared exclusively by *Xenoturbella* and Protobranchia (Fig. 2).

The oogenesis and nucleotide data are not fully conclusive by themselves but in combination they provide concordant morphological and molecular data showing that *Xenoturbella bocki* is neither a sister group of Bilateria nor from any other basal metazoan group but is a mollusc related to or within Protobranchia. This conclusion is drawn from apomorphies and not plesiomorphies and autapomorphies as earlier hypotheses. Further investigation, especially of embryology and biology, is needed for a complete understanding of *X. bocki*.

Olle Israelsson

Department of Zoology, University of Stockholm,
S-10691 Stockholm, Sweden
e-mail: olle.israelsson@nrm.se

Present address: Department of Invertebrate Zoology, Swedish Museum of Natural History, Box 50007, S-10405 Stockholm, Sweden.

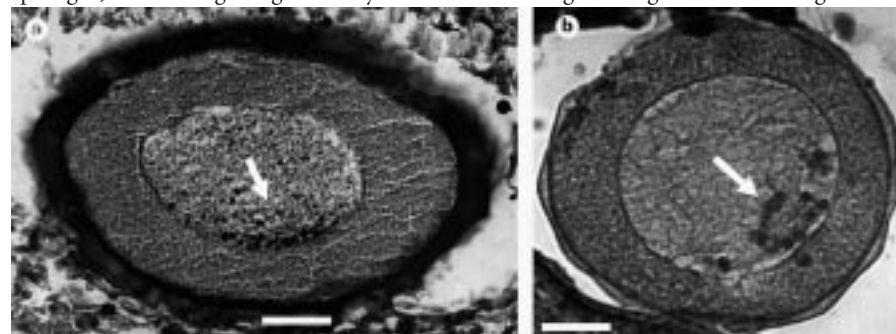


Figure 1 Relationship of *Xenoturbella bocki* to molluscs. Late vitellogenic oocytes of **a**, *Xenoturbella bocki* and **b**, the protobranch mollusc *Nucula nucleus*. Scale bars, 20 μ m.

1. Westblad, E. *Ark. Zool.* **1**, 3–29 (1949).
2. Ehlers, U. *Hydrobiologica* **227**, 263–271 (1991).
3. Norén, M. & Jondelius, U. *Nature* **390**, 31–32 (1997).
4. Jägersten, G. *Zool. Bidr. Upps.* **33**, 79–108 (1959).
5. Ehlers, U. & Sopott-Ehlers, B. *Zoomorphology* (in the press).
6. Reisinger, E. Z. *Wiss. Zool.* **164**, 188–198 (1960).
7. Franzén, A. & Afzelius, B. A. *Zool. Scr.* **16**, 9–17 (1987).
8. Pardos, F. *Acta Zool.* **69**, 1–12 (1988).
9. Pedersen, K. J. & Pedersen, L. R. *Acta Zool.* **67**, 103–113 (1986).
10. Franzén, A. *Zool. Bidr. Upps.* **31**, 355–482 (1956).
11. Pedersen, K. J. & Pedersen, L. R. *Acta Zool.* **69**, 231–246 (1988).
12. Giese, A. C. & Pearce, J. S. *Reproduction of Marine Invertebrates* (Academic, New York, 1975).
13. Harrison, F. W. *Microscopic Anatomy of Invertebrates* (Wiley, New York, 1991).

Discrimination down to a science

GATTACA

A film written and directed by Andrew Niccol

Columbia Pictures: 1997

Kevin Davies

Inside a darkened medical laboratory, the camera hovers over the shoulder of a young scientist as he stares intently at the fluorescing purple bands of DNA in the gel in front of him. Suddenly, he leans back and cries "Bingo!" A combination of *in vitro* fertilization and state-of-the-art DNA testing has yielded an eight-cell human embryo possessing the ideal combination of genetic markers — one that could not only spare the child-to-be from developing the fatal genetic illness running in its family but, even more remarkably, provide a lifeline for the family's oldest child who is certain to die unless a matching bone marrow donor is found.

No, this is not a scene from the new science-fiction thriller *GATTACA*, but a very real example from a recent television documentary of the power of modern medical technology, whereby techniques such as preimplantation genetic diagnosis can be used to select embryos untainted by one of at least a dozen fatal disease genes. The technique has produced more than a hundred healthy births so far (although, as in the case described above, embryo reimplantation frequently fails).

Now imagine this kind of technology carried to eugenic extremes, where the well-off can check off a host of desired traits while the embryo is still bathing in the Petri dish. In *GATTACA*, Andrew Niccol's disturbing vision of the not-too-distant future, the genetically elite 'valids' are a group of super-intelligent, buff specimens whose impressive 'genetic quotient' reflects the removal of known deleterious traits (aggressive tendencies, baldness) and the enhancement of others (usually height but, for the budding concert pianist, a pair of extra fingers). By contrast, naturally conceived 'in-valids' — their assorted neurological and developmental deficiencies diagnosed immediately at birth — are consigned to the genetic ghetto, uninsured and unemployed except for menial tasks.

Among them is Vincent (Ethan Hawke), whose congenital heart condition and low life-expectancy exclude him from joining the flourishing space programme. Undeterred, Vincent strikes a deal with Jerome (an excellent Jude Law), a crippled former athlete, to assume his 'valid' identity (becoming a 'de-gene-erate' to use the vernacular). As if orthopaedic surgery to add a few inches wasn't enough, Vincent's deceit must work at the DNA level, thus requiring him to carry Jerome's blood, urine, fingerprints — even



Splice girl: Uma Thurman would like to rearrange her DNA in *GATTACA*.

dandruff! Our hero pulls it off, but his rapid ascension up the ranks of the Gattaca Corporation, where he aspires to navigate the first manned mission to Titan, is suddenly jeopardized when the mission director is bludgeoned to death, leaving Vincent the prime suspect courtesy of an incriminating eyelash. As he dodges the police, Vincent seeks help in the form of his beautiful colleague Irene (Uma Thurman), who has a genetic make-up problem of her own to contend with.

GATTACA proclaims that "there is no gene for the human spirit", but the film will inevitably draw flak for fostering the erroneous impression that DNA and destiny are so closely intertwined. For example, recent studies clearly show that the influence of specific genes on personality traits such as 'novelty seeking' and 'neuroticism' is tangible but quite restricted, and that environmental cues are at least as important as the sum of the (mostly unknown) hereditary factors. The evidence points to genetic predisposition, but not predetermination.

In other respects, *GATTACA* is a surprisingly pedestrian affair given the controversial advertising campaign heralding its release ("Children made to order... it is now possible to engineer your offspring"). Niccol won't reveal how he thinks the gene selection process might work (no cute *Jurassic Park*-style cartoons here) or how one can obtain an instantaneous DNA readout from a single hair as easily as picking up a prescription (gene chips, perhaps?). What is more, Vincent's subterfuge would surely have been exposed if the intensive security collected buccal swabs instead of blood samples.

Nevertheless, *GATTACA* spotlights some thorny issues regarding genetic discrimina-

tion and technology in this, the tenth anniversary of the first funding for the Human Genome Project. Just how far might prospective parents be willing to go if offered the chance to shape their child's physical and psychological profile before birth? The exercising of parental choice in gender selection is hideously apparent in those parts of the world that practise infanticide. But in developed countries, an estimated 2,000 babies of preferred gender have been born by separating X- and Y-chromosome sperm before *in vitro* fertilization. Although mostly used to avoid sex-linked disorders, this service is also available for what one centre blandly terms "the purpose of family balancing".

Then there is the matter of intelligence. One clinic in California has produced hundreds of babies from the sperm of Nobel laureates despite criticism from the likes of the late Nikolaas Tinbergen, who observed that the narcissism of his fellow Nobelists "raised doubts about their possession of the very asset they want to pass on to their offspring: intelligence". Even so, James Watson probably got it right when he said: "If we could honestly promise young couples that we knew how to give them offspring with superior character, why should we assume they would decline?"

A final disappointment in *GATTACA* is the under-developed role of the magnificent Uma Thurman. An admiring former director of Ms Thurman said once: "She's like Garbo. Feature for feature it shouldn't work, but put them all together and you get this incredible face." There's a moral in there somewhere. □

Kevin Davies, former editor of *Nature Genetics*, is at the Howard Hughes Medical Institute, 4000 Jones Bridge Road, Chevy Chase, Maryland 20815-6789, USA.

The fragile contract

Endless Frontier: Vannevar Bush, Engineer of the American Century

by G. Pascal Zachary
Free Press: 1997. Pp. 518. \$32.50

Harvey Brooks

During the Second World War, Vannevar Bush, working with several like-minded colleagues, presided over a historic expansion of the role of organized innovation in society through the wartime Office of Scientific Research and Development (OSRD), which he helped to establish in 1940. The office was set up directly under the patronage of President Franklin Roosevelt, who gave it the authority and resources to initiate development of military technology without the prior approval or formal requirements of the military services.

Near the end of the war, Bush attempted to capitalize on his wartime experience by designing a postwar national science system, as outlined in his famous 1945 report "Science — The Endless Frontier", a classic document in science policy. Ironically, the organizational recommendations of the report, and even the content of the programme it advocated, were largely superseded as a result of other events. Yet the report did present a list of general principles to guide the relationship between the science and engineering community on the one hand and political and other social organizations on the other, principles which considered the two groups as equal partners rather than as servant and master. These principles have proved remarkably durable, helping to propel the United States to a position of scientific, technological, economic and military pre-eminence — a role model to which most of the world now aspires. They have left a legacy that many feel is destined to expand into more and more aspects of human life, although not without much opposition and questioning by those who perceive a conflict of these principles with the ideals of a true democratic society.

One indicator of this revolution is that, in every economically developed country, total national expenditures on research and development, both private and public, have grown from a fraction of a per cent of gross national product to nearly three per cent, while total national investment in innovation, defined more broadly to include production and marketing, although not well estimated, probably exceeds 10 per cent. Indeed, this statistic has come to be one of the common economic indicators that define the level of a country's economic and social development.

G. Pascal Zachary has written the first definitive and scholarly biography of Bush (1890–1974). It is based on a thorough review of Bush's newly accessible papers, informed by personal interviews with 54 of his contempo-



Bush: complex personality full of contradictions and inconsistencies.

raries. The work is extraordinarily well documented, with 73 pages of notes and a seven-page list of books and other sources.

The picture that emerges is of a complex personality with not a few contradictions and inconsistencies. Bush grew up and was socialized as an engineer in the classic mode of the late nineteenth and early twentieth centuries — that of the lone inventor and resourceful gadgeteer. It was a style that he never fully abandoned, and in which he often sought refuge from the frustrating world of bureaucratic politics, even while in the midst of creating a revolution both in the character of technological innovation as a product of large organizational teamwork, and in the role of experts in society. As a young professor at the Massachusetts Institute of Technology, Bush was an early pioneer in the design of analog computers. He thought about and published much on the potential of the computer and its capacity for amplifying the human intellect — a vision that was the acknowledged source of inspiration for many other early computer designers and entrepreneurs. Yet he was never comfortable with the digital age that eventually emerged and so was never part of the mainstream of computer development.

Bush was temperamentally a political conservative with an inherent distrust of government and politics and of large bureaucracies, both private and public. Yet at the same time he was a shrewd political operator who conceived of, and presided over, one of the largest and most successful networks of bureaucracies that the United States had ever seen. A compulsive solitary inventor who loved to potter in his own home laboratory and machine shop, he nevertheless realized and demonstrated the power of team research. An

advocate of competition, he energetically worked at eliminating 'wasteful' competition among the military services.

Initially sceptical of the possibility of an atomic bomb, he became a late convert on the advice of people he trusted and he pushed for the creation of the Manhattan Project outside his own domain of the OSRD. In the face of deep political scepticism, he was one of the first leaders to argue for internationalization of nuclear energy to head off an atomic arms race after the Second World War. He even favoured sharing the 'secret' with Stalin in the hope of averting such a race before it could start, accurately predicting how soon the Russians would test their own bomb.

Unlike the situation with the atomic bomb, Bush resisted through most of the Cold War any all-out effort for the development of an intercontinental ballistic missile or for manned space flight. Zachary points out this inconsistency, attributing it principally, but not entirely, to Bush's reversion to fiscal budgetary conservatism after the return to a more normal economy in the 1940s and to his growing scepticism about the inflated technological ambitions of the military that grew out of its wartime successes.

Zachary provides a good account of the origins of "Science — The Endless Frontier" and the vagaries surrounding its political reception in President Harry Truman's administration after Roosevelt's death. The report equated national welfare and military security with research strength, and outlined a system of permanent federal funding for research, mainly through contracts and grants, and mainly in universities and independent research institutes outside government. He also covers well the subsequent debates about the creation of the National Science Foundation, arguing that the sharp apparent differences among Bush, Harley Kilgore, John Steelman and Truman were much smaller and less consequential in practice for the future than the participants believed at the time, and probably did not justify the five-year delay in the eventual establishment of the foundation, an assessment with which I tend to agree.

A recurrent theme that runs through the biography and which was never resolved either by Bush, or, in my opinion, by the author, is the proper role of expertise and experts in a democracy. In the years just before the Second World War, Bush, along with many of his closest colleagues in a few élite universities and east-coast business communities, as they watched the German military build-up and later the Nazi war machine rolling over Europe, became concerned that dictatorships could run circles around democracies. After the war, he moderated this view, concluding that "all other things being equal, a democracy can outclass any despotism in bringing to bear on the struggle the combined efforts of science, industry and

military might". Yet he was convinced that it was only because of the war that democracy had performed so well. Remove the prod of war and the US government would revert to past form, as shown by the Truman administration's "exceedingly loose" handling of arms control in late 1945.

Bush continued to have a profound distrust of 'politics as usual' and aspired to achieving a nonpartisan expertise that could transcend politics and produce the right answer through realistic observation of the facts and analysis untainted by wishful thinking and political preconceptions. This aspiration was closely approximated in his experience with the OSRD. There, shielded from public view by military secrecy, and with unlimited access to Roosevelt, he could acquire the resources and authority needed to fulfil his vision of an independent civilian research-and-development organization that would serve as an equal partner with the military and the industrial leadership of the country, but which would be free to initiate developments without military blessing.

According to Zachary, Bush first conceived of this wartime organization in January 1940 — a remarkably bold vision. "Few civilians short of the president had ever imagined so grand a role in military affairs," says Zachary.

According to Zachary, "Bush himself later confessed that the creation of NDRC [National Defense Research Committee] was 'an end run, a grab by which a small company of scientists and engineers, acting outside established channels, got hold of the authority and money for the program of developing new weapons'. Such a power grab was necessary to launch a 'broad program... on an adequate scale', Bush insisted, but he conceded that it stoked resentment against him." Bush attempted similar tactics in getting Roosevelt to commission "Science — The Endless Frontier", but the strategy did not work out in the more open, participatory environment in which it was released after the end of the war.

Nevertheless, in Bush's mind there remained an inherent tension between the increasingly complicated and technical problems facing government, even in peacetime, and the country's democratic traditions. Bush's solution was a body of civilian technocrats, nonpartisan experts of the highest calibre, free of operational responsibilities for departments or agencies, with delegations of authority to make decisions in the name of the president. In Bush's view their effectiveness would have to depend as much on their detachment from politics as on their specialized knowledge.

This is in a sense the uneasy compromise at which the United States has arrived in practice today with the enormous growth of the staff of the Executive Office and a matching growth of congressional staff and support agencies. But, as Zachary remarks, "the specialization of expertise [has] advanced knowledge but also

expanded the possibilities for legitimate disputes about its import" with ample room for the intrusion of politics, especially in the overlapping domains between areas of expertise. We are still struggling for a solution today. This book does not provide one, but it does provide a wealth of new and valuable raw material for the continuation of the debate on expertise and democracy. □

Harvey Brooks is emeritus professor at the John F. Kennedy School of Government, Harvard University, 79 JFK Street, Cambridge, Massachusetts 02138, USA.

A bluffer's guide

Loose Ends

by Sydney Brenner

Current Biology: 1997. Pp. 128. £12.50, \$20

Walter Gratzer

In 1847, the aged Dr Routh, president of Magdalen and one of the foremost scholars of his day, was asked what precept he could offer to guide and sustain a young man embarking on a life of learning. After long thought, his face brightened and he spoke thus: "I think, sir, since you care for the advice of an old man sir, you will find it a very good practice *always to verify your references, sir.*" A pragmatic suggestion for our time might be "read *Molecular Biology of the Cell*, but don't drop it on your foot". This is very much the level of inspiration afforded by most scientific memoirs, often sponsored by charitable foundations, that have appeared in such numbers in recent years.

A collection of Sydney Brenner's writings arouses altogether higher expectations and I am happy to report that indeed it will disappoint no one, for here we have the authentic voice of the master himself. These pieces have appeared in the back of the monthly issues of *Current Biology*, which tend therefore to vanish from our library the day they arrive. The appearance of a new number is generally heralded for me by a call from a friend in a lab up the road, asking me whether I have yet read Brenner this month and regaling me with a few of the more outrageous squibs.

Loose Ends are reflections on biology and the scientific life. Uncle Syd's epistles to his nephew Willie treading the path of academic virtue from graduate student to institute director are laced with pungent anecdotes from an eventful life. Uncle Syd expatiates on the value of intellectual innocence and the treacherous nature of experimental phenomena; he informs Willy that the most abject of research students has the advantage of his professor, who has little time for such trivia as the work of his laboratory. "I have to warn you," he concludes, "that, sadly, this may be the only time in your career when you can enjoy research as an individual scientist."

As Willie comes to man's estate, Uncle Syd

proffers solutions to the problems that now face him — how, for instance, to avoid committees and conferences: the only acceptable excuse, he asserts, is to plead a prior meeting. To add conviction, Uncle Syd once invented a highly exclusive and mysterious society, which spawned numerous subcommittees to keep its members perpetually occupied. But at a pinch, he suggests, an inscrutable reply to an unwanted invitation, on the lines of "I regret that I am unable to accept your invitation as I find I cannot attend your meeting", will often serve, and has even been known to elicit a courteous acknowledgement. This is undoubtedly better than the telegraphic formula that Proust was said to have employed when he was being lionized by Parisian society: "Regret unable to come. Lie follows."

Willies the world over may also profit from Uncle Syd's tips on how to manipulate the old enemy, the bureaucrat. You can avert unpleasant decisions in committee by waiting until the administrators have formulated replies to your arguments and then confess that you were wrong after all and return to the original point. "You can do this," the wily Uncle Syd explains, "because the hallmark of a scientist is to be able to change one's views depending on the evidence; no administrator can do this." Here he is undoubtedly right. The legendary academic casuist Maurice Bowra gave it as principle never to allow scientists on committees: they were unreliable, he found, because their opinions could be changed by arguments. (It was also Bowra who announced, when outvoted by five to one, that the committee had evidently reached an impasse.)

Uncle Syd has also worked out how to use the telephone as a weapon, merely by reversing the postures of the caller and the respondent. This is an excellent device: when you have got past three secretaries and the Olympian grandee at the head of the organization finally comes on the line, you greet him with: "Why, hello, Sir Marmaduke, and what can I do for you today?"

In dilating on the seven deadly sins and their consequences, Uncle Syd again points a moral or two with some tantalizing anecdotes. Who was that editor of "an important biological journal" who submitted his work in an unworthy PhD thesis, examined and frivolously approved by Uncle Syd? The topic gives a clue, but a tormenting doubt hangs in the air. I believe, incidentally, that the seventh sin is not sloth, as here, but rather accidie, which lies somewhere between boredom and indifference. To me this is encapsulated by the perhaps rather sensible principle that if a research project is not worth doing at all it is not worth doing properly.

"Molecular Biology by Numbers" finds Brenner in top form with reflections on the science that he and a few friends mostly created. The number 1 stands for the principle of one gene, one enzyme, 2 the diploid chro-

mosome and the double helix, 3 the codon, 4 the canonical nucleotides and 5 the pentagonal faces in an icosahedral virus. Brenner's luminous insights repeatedly bring you up short as you read. There is also here a certain nostalgia for a glorious past, when the problems of molecular biology yielded to cogitation and to discussion in the coffee room, and not merely to a brutal experimental assault.

In those days theory was king, and Niels Bohr's principle that one should never believe an experiment that was not confirmed by theory held sway. Occam's razor was complemented by Brenner's noble conceit of Occam's broom, which was used to sweep out of sight the more inconvenient facts. Today's technology allows important facets of nature to be brought to light at such a rate that, says Brenner, "pausing to think about the results, or asking how cells really work, is likely to be seen as a source of irritating delay to the managerial classes, and could even endanger the career of the questioner".

He deplores also the advent of research-by-kit, for we seem indeed to be heading for a future in which not only the sequencing of DNA and the production of antibodies, but also the identification of new genes and their products and in due course the very formulation of the questions themselves, will be farmed out to commercial organizations for money. And then perhaps the bureaucrats will suddenly find Utopia within their grasp, because the uncouth and refractory scientists will no longer be needed.

Brenner is dismayed by the relentless advance of bureaucracy in general. We have entered an age of strategic mission statements and management training courses. (The UK Medical Research Council actually offers courses in how to appraise and discipline technicians, though of course they are not called that now.) Security is another of his bugbears, although he does not mention safety, which is now the prime growth area in many organizations. To my mind, much the most dangerous place in most laboratories is the only one the safety officer habitually spares — the library, with its ever-present hazard that a bound volume of the *J. Biol. Chem.* will fall from a high shelf and make an end of you.

If, then, you are not the person who has razored out the last page of *Current Biology* and you want a guide to how to comport yourself in your scientific career and to survive and even thrive, then I urge you, for your pleasure and profit, to buy Uncle Syd's book. You may possibly (especially if you are one of the camp-followers of the profession) want to gratify him with the abusive letters and obscene telephone calls of which he has so far been disappointed. As the old bruiser Kingsley Amis once observed: "If you can't annoy somebody with what you write I think there is little point in writing." □

Walter Gatzert is at the Randall Institute, King's College London, London WC2B 5RL, UK.

Life's dead letter

Lifelines: Biology, Freedom, Determination

by Steven Rose

Allen Lane: 1997. Pp. 335. £20. To be published in the United States by Oxford University Press in January at \$30

Benno Müller-Hill

Genetics has been one of the most successful sciences of the twentieth century. Yet its history is not a simple success story. In the Soviet Union, genetics was almost destroyed. In Germany, it served as an argument for mass murder. In the United States, it was used to legitimize mass sterilization, laws against racial mixing and a restrictive immigration policy.

And today, a sizeable minority of the public looks at developments in genetics with disrespect.

Steven Rose is a biologist who has claimed in many books that something is deeply wrong with biology itself. In his latest offering, he suspects that the culprits are reductionism and biological determinism. He writes: "The challenge of the opponents of biological determinism is that, while we may have been effective in our critique of its reductionist claims, we have failed to offer a coherent alternative framework within which to interpret living processes.... *Lifelines* originated as an attempt to meet the challenge."

Rose begins by presenting a brief history of biology and genetics. In a discussion of T. H. Morgan's work on mutant fruitflies, he writes that "they had red rather than white eyes". It was in fact the other way around! He then points out that "the proportion of unusual characters in the fly population could be greatly increased by stressing it in some way — for instance by exposing the flies to not-quite toxic concentrations of particular chemicals, or to radiation such as X-rays". It was H. J. Muller, not Morgan, who discovered the mutagenic action of X-rays. The reader will be confused by Rose's account of history.

One of Rose's central points is that genetics, which deals with single genes and gene products, is unable to illuminate the processes of development and evolution: "The great expansion of genetic knowledge in recent decades has yet to be matched by a comparable increase in the understanding of development."

He neglects to mention the work of Ed Lewis or Christiane Nüsslein-Volhard and their colleagues. In fruitflies, the first stages of development are now well understood, and in fish we may know them soon.

It is true that mouse and human development are still mysteries. But these mysteries are challenges to geneticists. Rose sees it dif-

ferently: "Reductionist ideology has a number of serious consequences. It hinders us biologists from thinking adequately about the phenomena we wish to understand." I think he is wrong, and I hope he is unsuccessful in selling his idea to young biologists who might read this book.

Rose then sets out to attack the 'ultra-Darwinists' Daniel Dennett and Richard Dawkins. Again, I think Rose's point is badly taken. "To isolate from this evolving web a single factor, be it a gene or organism, as the unique determinant of change is as problematic as isolating a single enzyme from the metabolic web that constitutes the cell. Any such attempt at isolation is a reductionism that mistakes method for theory." I and many others have put some effort into isolating such genes and I remain convinced that the results of our studies have better illuminated our understanding of the living cell.

After warning his readers against the dangerously seductive character of metaphors, Rose presents us with those of his own choice.

Here is a selection: "We must speak of the dialectic of specificity and plasticity during development, the dialectic through which the living organism constructs itself"; "...to offer a perspective on biology which transcends genetic reductionism, by placing the organism, rather than the gene at the centre of life — this perspective I call homeodynamic"; "Our lives form a developmental trajectory, or lifeline, stabilised by the operation of homeodynamic principles. This trajectory is not determined by our genes, nor partitioned into neatly dichotomous categories called nature and nurture. Rather it is an autopoietic process". These quotations give the flavour of what is to replace reductionism if we follow Rose.

So we come to what seems to me the central issue. Rose argues that "the argument against hunting for neurogenetic explanations (for violence etc.) is not that it is immoral or unethical to do so. It is simply that, despite the seductive power of reductionism it is the wrong level of the disciplinary pyramid". To my mind, violent crime is not part of science. If it is made part of genetics, it converts genetics into an ideology or religion, as has been demonstrated in Nazi Germany.

Finally, Rose sets about "making Biology whole again" by delivering, as a new Moses of biology, "biology's Decalogue". As far as I could tell, it bears no resemblance whatsoever to the old Decalogue. To quote its last sentence: "And it is therefore our biology that makes us free."

If I had to choose, I'd prefer the Mosaic version. □

Benno Müller-Hill is at the Institute of Genetics, University of Cologne, Cologne 50931, Germany.

Après lui le déluge

Buffon

by Jacques Roger

Cornell University Press: 1997. Pp. 492.

\$49.95, £39.50

Philippe Janvier

When about to be guillotined in 1794, "Buffonet", the unfortunate son of the French naturalist Georges Louis Leclerc, Count of Buffon (1707–88), appealed to the people by shouting "Citizens, my name is Buffon!"; in the hope that this would perhaps save him. He was aware that the considerable fame of his late father, who had died a year before the French Revolution, had survived the turmoils of the terror.

Buffonet failed, but Buffon still remains dear to the heart of the French layman. Until the 1950s, French popular books on animals frequently referred to Buffon (*The Children's Buffon*, *The Family's Buffon*, and so on).

What made Buffon so different from other 'savants' of the past two centuries? He was admittedly a pioneer of many new fields, tackling such issues as the antiquity of the Earth, the biological species concept and historical biogeography. But time has passed; we now live in the age of evolution, dominated by the glorious shadow of Darwin, and Buffon's writings sound obsolete.

Jacques Roger (1920–90), a world-renowned historian of eighteenth-century science, devoted most of his academic life to Buffon. His book, first published in French (*Buffon, un Philosophe au Jardin du Roi*, 1989), is now translated into English — admirably so — and provides not only a biography of Buffon, but also a thorough analysis of his views in the framework of the philosophical debates in the second half of the eighteenth century.

Roger points out previously unnoticed aspects of Buffon's innovative insights, but he also puts back in their right place several overinterpreted sections of Buffon's 44-volume *Histoire Naturelle Générale et Particulière*.

The French are usually prone to defending humble scientists who are victims of oblivion or authority, such as Lamarck, but their sympathy for Buffon is at odds with this tradition. Buffon was not a victim, nor a rejected genius, nor an introverted scientist. He loved life, food, women, wealth, power and honours, and he was perfectly conscious of the success of his popular style. He was also a good businessman and a talented manager.

Nevertheless, he was moderate enough in both his scientific statements and his political involvements to avoid having severe problems with either the censor or the progressive circles of his time. He seemed to cope very well with his society, although he fore-

saw the revolution ("I see a large movement coming, and no one to lead it").

What the layman liked above all in Buffon was his skill for popularizing the natural sciences, largely using common sense as a basis for deeper questions, covering all fields from astronomy to anthropology, and all this in an explanatory historical framework that did not seem too shocking to religious minds.

The situation was quite different in academic circles. During his lifetime, Buffon faced strong criticism from some French and foreign scientists (in particular the followers of Linnaeus) who accused him of being a superficial observer, a flabby experimenter and an "empty and bombastic" thinker. During the revolution and empire, Buffon faded into the past as a typical character of the *ancien régime*, although his fame remained great among the public, as shown by the numerous editions of his *Histoire Naturelle*.

With Darwin and the rise of modern evolutionary thought, and until recently, some tried to find in Buffon's works cryptic allusions to descent with modification. The aim was to show that Buffon was a pioneer of this idea as well, rather than Lamarck who, unjustly, bore alone the burden of the theory of the inheritance of acquired characters. The famous chapter on the ass, where Buffon precisely describes all the evolutionary consequences of the classifications and suddenly rejects them as if in fear of the censor, is often cited as evidence for his transformist convictions.

Roger, however, rejects this interpretation. When Buffon feared the censor, such as in the case of the age of the Earth, he used to express his ideas in a cautious way, as a mere theory, but with numerous factual arguments in support of it. On the contrary, in the case of the 'transmutation' of species, his rejection was straightforward because his strict species concept, based on the immutability of the "interior mould" and intersterility through time, could in no way allow it.

Moreover, had he been a transformist, he could have referred to Maupertuis who, in 1751, had already proposed a theory of evolution based on successive, fortuitous mutations within species (to which Darwin later added natural selection). But Buffon rejected this view with a number of arguments that would be repeated by antievolutionists throughout the nineteenth century.

Chapter 14 of Roger's book will be extremely useful to historians of biology, as it clearly explains Buffon's concepts of species, genus and family, which are different from Linnaeus's view and, in addition, changed with time.

Nevertheless, Buffon's ideas about species and systematics are strangely convergent with those of some modern systematists and phylogeneticists. He considered the species as the only entity in nature that sur-



"Courly rouge du Brésil," from Buffon's *Histoire Naturelle des Oiseaux*.

vives unchanged in time, and he regarded groups of species (taxa) as mere speculations based on arbitrary sets of characters. Modern systematists would add that these sets of characters no longer need to be arbitrary if demonstrably inherited from a common ancestor and, in a sense, are more real than species themselves, but they would agree that taxa are not real.

This book is full of amusing anecdotes, such as the vivid reaction of Thomas Jefferson to Buffon's statement that North American animals were smaller than European ones because North America was cooler than Europe at equal latitude (Buffon was obsessed by the relationship between heat and life). This was unacceptable for Americans and later led Jefferson to try to find living mammoths in the West.

Roger's book reads like a novel but also explains the rise of Buffon's ideas in the general context of the debates of his time. Its numerous notes, references and extensive index make it a major tool for historians of science. I would also recommend it to professional biologists and biology students, for whom serious biology often starts only with Darwin.

The book clearly shows that Buffon addressed many questions that are still pertinent to biology and that some sections of his *Histoire Naturelle* are amazingly modern in comparison with later, nineteenth-century views (in particular his observations on the uniqueness and biogeographical history of the human species).

The English translation by Sarah L. Bonnefoi is superb, as it preserves both the accuracy of Roger's style and the elegance and subtlety of Buffon's, which he himself regarded as most important in conveying his ideas. □

Philippe Janvier is at the Laboratoire de Paléontologie, Muséum National d'Histoire Naturelle, 8 Rue Buffon, 75005, Paris, France.

Students rule

Academic Duty

by Donald Kennedy

Harvard University Press: 1997. Pp. 310.

\$29.95, £19.95

Daniel S. Greenberg

Donald Kennedy, president of Stanford University from 1980 to 1992, sets out to prescribe for the many, long-lamented ills of academic institutions — neglect of teaching, exploitation of graduate students, discouragement of women science students, inordinate glory-seeking, and many more. But, along the way, he seems to suggest that big academic institutions are rigid, arrogant and powerful, and not likely to mend their ways.

The “academic duty” of the title refers to the obligations of teaching, mentoring, truthfulness and community service that he urges upon the professoriate. But this remarkable book, sure to stir debate in and beyond academic circles, might just as well have been titled *Requiem for Academe*. Whereas industry, government and other

sectors of society have responded to major economic and social changes, the big universities constitute “a portrait of conservatism, perhaps even of senescence”, he concludes.

The so-called research universities, the fount of some 60 per cent of doctoral degrees in the United States, “must be the agents of change”, because they set the standards and tone for much of higher education, Kennedy claims. But, recognizing a strong spirit of innovation in less-renowned colleges and universities, he observes: “The difficulty is that [the major institutions] are both successful and prestigious, and they lack the natural appetite for renovation and reform that characterizes the striving, transformational institutions. But unless they change, little else will.”

The problem is that the system tends to defy change, he says. Confronted with a PhD glut, “the most prestigious institutions invariably argue that it is other, less excellent institutions that should cut back production by limiting admissions”. Despite an oversupply in the biomedical fields, he points out, the National Research Council in 1994 blithely recommended maintenance of predoctoral awards for basic science at the 1993 levels. Graduate students provide inexpensive, obedient help for their tenured masters. Does this figure in the unwillingness to curb their numbers? Kennedy does not contest the possibility.

He illustrates his argument with scores of episodes derived from his long career as a highly productive biologist, commissioner of the US Food and Drug Administration during the Carter administration and university president. Symbolizing the academic intransigence that he deplores is a tale told by the chemist Carl Djerassi, a Stanford colleague. Noting an absence of mentoring in his department, Djerassi proposed a questionnaire for graduate students and postdoctoral fellows on the guidance they received from their lab chiefs on such matters as record keeping and publication policy. “The departmental faculty voted not to permit the questionnaire to be circulated”, Kennedy reports. “Plainly, these subjects are uncomfortable ones. Given the faculty’s reluctance to talk about them, it is small wonder that generation after generation, we launch innocents into the world of academic mentorship.”

As often before, educational reform “is again at the centre of concerned conversation in colleges and universities”, Kennedy says. But “the results are still meagre”.

Academic Duty originated in Kennedy’s concern about the failure of academic institutions to prepare aspiring academics for their responsibilities as members of a university community. “My own aim,” he explains, “is to write primarily at, about, for members of the faculty: their central role in

the institution’s mission, the way they relate to their legal owners and managers, and their responsibility to students.”

On this last point, Kennedy is uncompromising, and extravagantly optimistic about a panacea, arguing that “improvement must entail putting students and their needs first. Once that is done, the rest falls into place: the complex challenges posed by intellectual property disputes, the tension between teaching and research, the ethical problems in faculty–student relationships, professional misconduct issues, the need for creative thinking about undergraduate education reforms — indeed, all the manifold difficulties so prominent in the growing public distrust of our academic institutions. Putting students first is a simple design principle, but it has great power.”

There can be no argument about the value of putting students first. But how this would reduce, let alone eliminate, scientific misconduct and squabbles about intellectual credits is not apparent or explained.

After his service in Washington as commissioner of the Food and Drug Administration from 1977 to 1979, Kennedy began a long, successful run as president of Stanford. His resignation followed a collision in 1991 with sulphurous Congressman John Dingell, on the warpath in quest of illicit academic gorging on federal funds. *Academic Duty* could easily have excluded the grisly encounter that ensued when Dingell publicly lacerated Kennedy and Stanford for allegedly using federal research funds for entertaining, personal effects for Stanford’s presidential residence, and other purchases bound to raise taxpayers’ ire. But, in abbreviated form, Kennedy tells the complex story, emphasizing that the university was eventually exonerated of any wrongdoing, and attributing the blow-up with Dingell to the federal government’s inscrutable regulations for calculating overhead costs on research grants. “Certainly,” Kennedy concedes, “the government shouldn’t pay 23 per cent of the cost of flowers for university entertaining.”

The relevance of this episode to academic duty is distant and might have been saved for a fuller treatment by Kennedy elsewhere. The same applies to his quick reviews of several prominent cases of scientific misconduct, notably the highly publicized Gallo and so-called Baltimore cases. In taking up the issue of academic duty, Kennedy has performed a valuable service directed at rescuing our great research universities from their entrenched follies. At various points he expresses guarded optimism, but the evidence he offers does not justify optimism. □

Daniel S. Greenberg, founding editor of *Science & Government Report*, is at 3736 Kanawha Street NW, Washington DC 20015, USA.



Kennedy: prescription for fitter research universities.

Credit control

Ronald Ross: Malariologist and Polymath – A Biography

by Edwin R. Nye and Mary E. Gibson
Macmillan/St Martin's: 1997. Pp. 316. £45, \$59.95

Mary J. Dobson

On 20 August 1897, Ronald Ross made the critical observations that were to prove that mosquitoes transmit malaria. On that day, in Secunderabad, India, Ross peered down his cracked microscope and saw pigmented swellings (oocysts) embedded in the stomach wall of an *Anopheles* mosquito that had been fed on a malarious patient. In the stifling atmosphere of a dingy little military hospital, and unable to use the *punka* (ceiling fan) for fear that his dissected mosquitoes might blow away, Ross recorded his first signs of success.

After several years of frustration, while working on the problem in the Indian Medical Service, he had at last established that malaria parasites, identified by Alphonse Laveran in 1880, spent part of their life cycle in 'dappled-winged' (*Anopheles*) mosquitoes. The mists of malaria were lifted and the following day Ross celebrated his discovery in verse that included the moving lines: "I find thy cunning seeds, O million-murdering Death".

Ross's breakthrough has often been cited as one of the most dramatic discoveries in medical science, and Ross himself was anxious to remind the world of its significance and to reap its rewards. It is a fitting tribute to Ross that Edwin R. Nye and Mary E. Gibson have produced this elegant and illuminating biography during his discovery's centenary year. Their book is more than an account of a great discovery. It reveals Ross (1857–1932) as a family man, malariologist and polymath who was caring, committed and, above all, determined to receive due recognition and financial reward for his discovery.

Ross emerges from this biography as one of the most active, ambitious and embittered pioneers in the history of tropical medicine. We read of his personal and scientific activities, extensive travels, interactions and quarrels with scientists, politicians and writers, his time at the Liverpool School of Hygiene, his literary and mathematical achievements, and the many accolades he received, including a fellowship of the Royal Society, the Nobel prize for medicine in 1902, a knighthood, and the legacy of the Ross Institute.

The authors have no shortage of material from which to construct their narrative. Much of Ross's story is already known from his *Memoirs*, which won him the James Tait Black Memorial Prize for biography in 1923 and were republished in part in 1988 as *The*



From life cycles to tricycles: Sir Ronald and Lady Ross, 1929.

Great Malaria Problem and its Solution. Ross's self-portrait is richly reinforced and supplemented by Nye and Gibson's selection of extracts from the unpublished Ross archives.

Ross was so obsessed with the need to keep a record of his personal achievements and to prove his priority in solving the great malaria problem that he requested and managed to obtain copies of letters he had sent to his numerous correspondents.

It is through his writings that the authors describe Ross's bitter disputes with colleagues and authorities, particularly his rift with Patrick Manson, the 'father of tropical medicine' who had first encouraged Ross to explore the link between malaria and mosquitoes, and his venom towards the Italian malariologist Giovanni Grassi. It would, however, have been interesting to have included more about the scientific activities and contributions of the Italian malariologists. For although Ross, while stationed in Calcutta in 1898, was able to follow up his initial discovery by showing the life cycle of bird malaria, it was the Italians who, later in the same year, worked out the complete life cycle of human malaria. We will, perhaps, have to await the Grassi centenary in 1998 to get the other side of the story.

The Ross archives played an important role in Ross's political manoeuvrings. In 1928, having unsuccessfully petitioned the British Parliament to reward his research, as it had Edward Jenner a century earlier when he discovered the smallpox vaccine, Ross made his grievance public by selling his personal papers for £2,000, a large sum at the time. He also fuelled a lively correspondence in the scientific literature, entitled "The Rewards of Research and How to Apply For

It". Undoubtedly, this dissatisfied scientist would have been well-satisfied that his biographers have given prominence to his strong personal feelings of pecuniary injustice—an issue as troublesome to scientists today as it was a century ago.

The rewards of Ross's research must, however, be seen in another light. Ross, Grassi and others not only transformed the understanding of malaria but they laid the foundations for the application of work on the practical control of the disease. For the first time in history, it was possible to show that, if the mosquito vector could be eliminated, malaria transmission could be prevented.

Shortly after his discovery, Ross turned his attention with passion and zeal to the prevention of malaria. Although the authors make an excellent job of portraying Ross as malariologist and polymath, they are much less effective in discussing the early schemes of malaria control that were opened up by Ross's pioneering work. They also miss an opportunity to examine Ross's ideas for the prevention of malaria, including the cost, value and feasibility of 'mosquito brigades', chemoprophylaxis and bed-nets, in the context of current policy.

The successes and failures of anti-larvae measures immediately following Ross's discovery, especially the contrasting outcomes in Panama and at Mian Mir in what is now Pakistan, not only demonstrated the possibilities and limits of eliminating the mosquito vector but also set in motion a series of dichotomies about how best to control malaria. The legacies of these early debates, which are not developed in this biography, have resurfaced throughout the twentieth century and are still with us. Indeed, it is a sad irony that, 100 years after Ross's brilliant dis-

covery, ways of tackling the great malaria problem remain unresolved.

At the Second Global Meeting of Parasitology held in Hyderabad, India, in August to commemorate the centenary of Ross's discovery, the hope, the frustration and the determination to control malaria were as much in evidence as they had been in the early twentieth century. If one lesson has to be learnt from this history, it is the difficulties of moving forward from the results of scientific discovery.

Ross would have been delighted by the attention he has received in this centenary year. Not only have we welcomed this biography, we have also witnessed the renaming of a road and a hospital in his honour, the renovation of his laboratory in Secunderabad, the unveiling of Ross busts, and his picture projected on postage stamps and T-shirts. But Ross would not have been happy to learn that 300–500 million people are still suffering from malaria 100 years after 'Mosquito Day' 1897.

I hope that this biography will be read both for its own merits and as a way of stimulating further interest in Ross's views and vision for the prevention of malaria. □

Mary J. Dobson is at the Wellcome Unit for the History of Medicine, 45–47 Banbury Road, Oxford OX2 6PE, UK.

ADVERTISEMENT

"How wonderful that Jacques Roger's masterful study of Buffon and his work is now available in an admirable translation. . . . The perfect example of what a fine biography should be. [Roger] not only presents Buffon as a person, he also gives a most illuminating critical analysis of Buffon's thoughts, beliefs, theories and controversies."

—ERNST MAYR,
Cornell University Press
Book Review

JACQUES ROGER

Buffon

A Life in Natural History Translated by Sarah Lucille Bonnefoi

16 illus. \$49.95
At bookstores, or (800) 666-2211

CORNELL UNIVERSITY PRESS

READER ENQUIRY NO. 578

Heavyweight history

Science in the Twentieth Century

edited by John Krige and Dominique Pestre
Harwood Academic: 1997. Pp. 941. £80, \$120

Graham Farmelo

One of the saddest sights in science is the chasm between its practitioners and its historians. Although scientists are curious about the history of their subject and most historians of science have an interest in contemporary science, the professions have yet to establish an effective dialogue.

This cultural division was exemplified two years ago when the British and American Institutes of Physics published a three-volume history of twentieth-century physics that featured contributions from 31 distinguished scientists, not one of them a trained historian. The result was an impressive but introspective account that would have been a good deal richer if the contributors had brought to their essays a wider historical awareness.

Science in the Twentieth Century goes to the other extreme: its contributors are 50 leading historians, none of whom is a practising scientist. At first sight, its 941-page bulk and the austerity of its design make it one of those books that cry out to be weighed rather than read.

But a cursory look through its pages reveals it to be something rather special. For these are authoritative essays intended for scientists and lay people with an informed interest in modern science. Almost all the contributions are devoid of historiographical jargon and have relatively few references, so the authors are not writing solely for their peers. The book, or "project" as its editors tellingly describe it, is an ambitious and laudable attempt to make the best thinking about the history of modern science accessible to scientists.

There is no denying the sheer range of the material covered. Between them, the authors review the nature of science, its place in the social fabric, how the most productive research has been carried out and what has been achieved. Although the content of the essays reflects the pre-eminence of the scientific contributions from the United States and Europe, it is good to see thorough reviews of the work done in other parts of the world, notably Russia, India, Latin America and Japan.

Similarly, the branches of science considered go way beyond the historians' favourite areas of medicine, astronomy and fundamental physics; between these covers are first-rate histories of disciplines as diverse as geology, mathematics and polymer chemistry. To read this book is to begin to appreciate how much our understanding of science (not to mention technology and industry) has been enriched by the comparatively small community of historians of science since their subject became

established as an academic field, only about 50 years ago.

Invidious though it is to focus on a few of the 46 essays in this tome, three will serve to give an idea of the breadth of coverage. In one of the three introductory pieces, Simon Schaffer of the University of Cambridge considers how science has been perceived by the public in different countries during the past century. In a characteristically lucid piece, he examines a wide variety of media and comes to the surprising conclusion that the most important way in which the public has derived its image of science is through exhibitions. (It would be fascinating to use survey data to check whether this hypothesis holds in the future.)

In this most violent of times in human history, every branch of culture has been influenced by the military, and science is no exception. In a masterly review, Everett Mendelsohn of Harvard University looks at the impact on Western science of the armed forces, notably in the world wars. He points out that after the First World War none of the major powers developed a clear policy for science in the military, with the result that on the eve of the Second World War most of the agencies linking science and industry had to be reinvented. Now, in the aftermath of the Cold War, Mendelsohn suggests that scientists have the "illusion of autonomy", a notion with which he abruptly and tantalizingly ends his piece.

One branch of science served particularly well in the book is immunology. Anne Marie Moulin, of L'Hôpital des Enfants Malades in Paris, reminds us that this subject is not much older than the century itself. She gives a fascinating account of its history, which she believes is best characterized as a series of controversies rather than of Kuhnian revolutions. How much has been achieved in these disputes will be tested, Moulin asserts, in the "clear-cut assignment" of tackling the AIDS epidemic.

This book would be an asset to any science library. For scientists who have the time to read it and the stamina to carry it back to their offices, there is much to be learned and enjoyed. They may be disappointed that the index does not do justice to the richness of the text and that the editors do not provide a coherent summary in their otherwise admirable introduction. Also, many will regret that there are no contributions from historically aware scientists, who could have given a more vivid sense of what actually happens at the coal-face of research.

There is, however, no doubt that the editors deserve our gratitude for all they have done, as we approach the end of science's most productive century, to bring us this monumental collection. It is of much more than historical interest. □

Graham Farmelo is at the Science Museum, Exhibition Road, London SW7 2DD, UK.

Inner visions

The Ovary of Eve: Eggs and Sperm and Preformation

by Clara Pinto-Correia
University of Chicago Press: 1997.
Pp. 376. \$29.95

Frank Gonzalez-Crussi

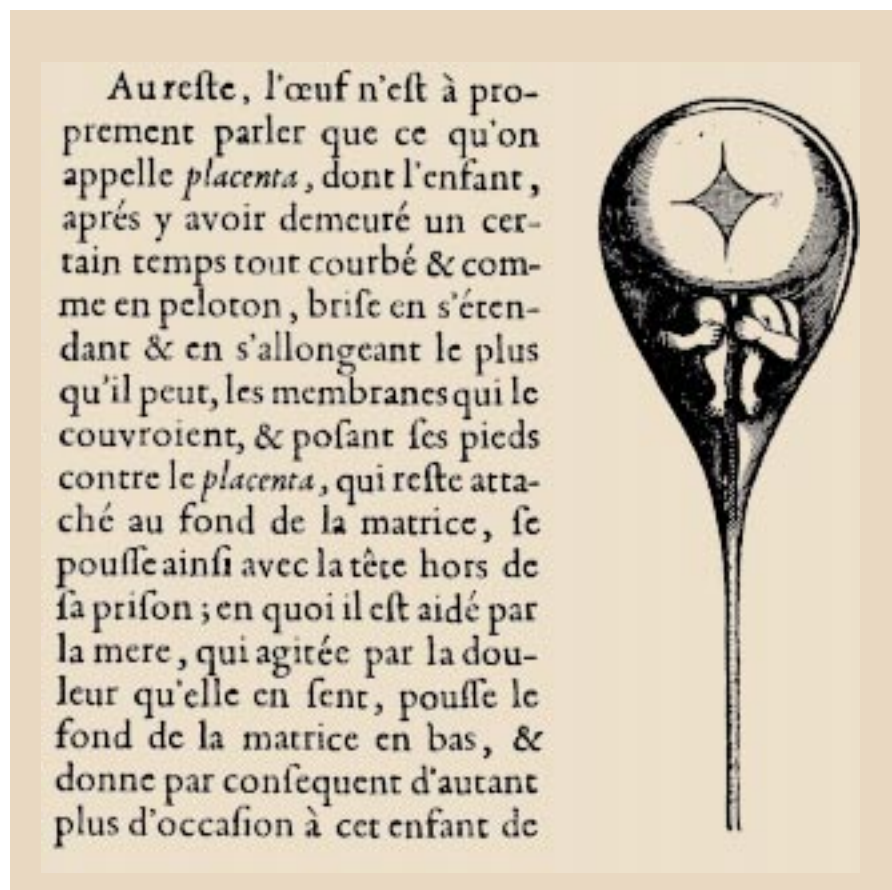
How do we come into being? Few questions are as fundamental, or as appealing to the imagination, as this one. The process by which we are shaped into existence is a phenomenon at once commonplace and portentous. Millions are conceived and millions born every day, following an unvarying plan. Yet who could fail to be filled with wonder when confronted by the spectacle of an inchoate, tiny germ that develops symmetry, forms eyes, sprouts limbs and ultimately becomes a fully formed individual?

Considering the inexhaustible suggestiveness of this process, and the amount of folly, passion, fantasy, wild speculation and serious investigative effort that it has given rise to, it is surprising that relatively few readable historical accounts have been written. True, histories of embryology have been compiled before, but most fail to do justice to the multi-hued tints and ramifications that their subject matter offers at every juncture. This deficiency is amply corrected by Clara Pinto-Correia.

Her engrossing book is likely to remain for many years the most readable and informative — certainly the most enjoyable — account of the early scientific efforts to explain human conception. This is not surprising. As a professor of developmental biology, the author has the necessary background to assess the relative importance of the various contributions; as an established novelist and poet, she brings to the task clear, agile prose, and a gifted eye for vivid, evocative and memorable material.

When old ideas and hypotheses, no matter how exalted, are superseded by new evidence, they must be cast aside. This attitude is inherent in the much-revered scientific method, and is no doubt commendable in the scientist. The historian, however, ought to prize the obsolete. Yet, when describing discarded hypotheses, historians often sound steeped in the conviction that they are merely chronicling past errors. Their tone tends to turn contemptuous, and the reader senses their amazement at the credulity of our forebears, by whom the absurd, far-fetched notions were sustained.

It is a strength of *The Ovary of Eve* that it succeeds in fully appropriating the viewpoint of the scientists of the past, and makes the reader realize the nobility, the earnestness, the inevitability and the beauty of many of their constructs, even those condescend-



Human fetus in a sperm from Nicolas Hartsoeker's *Essay de Dioptrique* (1694), which is largely a treatise on optics and optical instruments. His extreme preformation theory proposed that "the first males would have been created with all those... that would be born until the end of the centuries".

ingly dismissed as 'foolishness' by less perceptive historians. "It is a sad fact," the author tells us in the prologue, "that history is written by the victors. In this book, I want to challenge this trend, and tell the marvellous story of the losers."

'Marvellous' is right, and rendered more so by the lively style of the narrative. But, as the book comes to an end, it is by no means clear that there are winners and losers. The reader is made to realize that the confrontation between preformationism and epigenesis is far from over; that the fundamental concepts remain the same, even though the warring factions take on a different guise each time; and that we are not quite through with mulling the notion, first propounded early in the present century, that ontogenesis is "a cytoplasmic epigenesis, underlain by a nuclear preformationism".

Or, as Stephen Jay Gould asserts in the foreword, our present view of morphogenesis is epigenetic but, to the extent that we acknowledge a pre-existing 'blueprint' that directs all activities, preformationism may still be said to be alive and well and domiciled in the nucleus.

The subtitle, *Eggs and Sperm and Preformation*, is inaccurate. The book is much broader in scope and richer than that. We see Lazzaro Spallanzani fitting frogs with boxer

shorts to ascertain the role of semen in fecundation (but to the end believing that sperm cells were mere contaminants). We find the microscopists — Leeuwenhoek, Hartsoeker and Du Bois-Regard — expounding their ideas and flinging sarcasm at each other. And whereas Ambroise Paré and Ulisse Aldrovandi explain the causation of monstrous births, Vandermonde, eugenicist *avant la lettre*, enlightens us on the ways to perfect the human species.

But there is much more. Reference to Hartsoeker's 'homunculus' leads to considerations on the Egyptian goddess Ka, the work of the Arab alchemist Jābir Ibn Ḥāyan, the Jewish myth of the Golem, Francesco Redi's experiments on generation, the fairy tales of François de la Plantade and so on. Interesting and erudite connections abound in each of the eight engrossing chapters.

The book is an individual and idiosyncratic work. That may be one of the reasons why it accomplishes so well the ends that Vandermonde, according to the prologue, wished for his own production: "To first detain the reader before novel and amusing objects, that in treading the path of beauty he may arrive effortlessly to that of science." □
Frank Gonzalez-Crussi is in the Department of Pathology, Children's Memorial Hospital, Chicago, Illinois 60614, USA.

No direction known

The End of Certainty: Time, Chaos, and the New Laws of Nature

by Ilya Prigogine

Free Press/Simon and Schuster: 1997. Pp. 228. \$24, £20

In Search of Lost Time

by Derek York

Institute of Physics Publishing: 1997. Pp. 141. £7.95, \$15 (pbk)

Huw Price

Can physics explain the difference between past and future? The problem is that the laws of physics seem to be time-symmetric: if they allow a process with one temporal orientation, they allow it in reverse. Yet many ordinary processes seem to be irreversible. Ilya Prigogine calls this the 'time paradox', and argues that the solution lies in chaos theory and related methods pioneered by himself and his Brussels colleagues — a radical alternative, he believes, to a tradition due to Ludwig Boltzmann.

Prigogine paints with a broad brush. For example, he seems to equate explaining the thermodynamic asymmetry with finding a physical basis for the apparent flow of time. But these are different issues. Someone who treats time on a par with space, denying that there is objective flow, will still acknowledge that the Universe is asymmetric along its tem-

poral axis, in the way described by the second law of thermodynamics. If Prigogine's methods do explain thermodynamic irreversibility, they simply account for this asymmetry — they do not give us objective flow.

Leaving flow to one side, what is it that actually needs to be explained about the asymmetry of thermodynamics? Ordinarily we say it is the fact that entropy increases. But whether entropy is increasing or decreasing depends on which we take to be the 'positive' direction of time. Reverse the scale on the graph, and an increase becomes a decrease.

Is one choice of scale objectively the right one? Is what we call the future objectively the sense of 'positive' time? If so, this objective fact must consist of something other than the thermodynamic asymmetry itself. Otherwise, asking why entropy increases is like asking why a snake's head is at the front end: it would be true more or less by definition.

So there is a dilemma for Prigogine. Either he needs to pull two rabbits out of the hat — one to account for the objective direction of time itself, and another to account for the thermodynamic asymmetry — or the puzzle about irreversibility is not what he thinks it is. It isn't that entropy increases *per se*, but rather that it increases in one direction and decreases in the other. (Analogously, the real puzzle about snakes is that the two ends are so different.)

In the latter case, Prigogine's methods are surely being applied to the wrong task. He suggests that the sensitivity of chaotic systems to initial conditions introduces unpredictability into physics, and hence explains why entropy is bound to increase. But the sensitivity applies whether we start with what we call initial conditions and work 'forwards' or start with final conditions and work 'backwards'. The more chaos helps to convince us that entropy increase is inevitable in one direction, the more it ought to amaze us that we see decrease in the other direction (a fact that surely ought to give us pause about the reliability of the 'explanation' in the former case: how good is a method that gets things so badly wrong in one case out of a possible two?).

Boltzmann came this way in 1877, noting that it was a consequence of his new statistical methods that entropy ought to be higher in the past as well as in the future. The real 'time paradox' is that entropy goes down in one direction, not that it goes up in the other. Boltzmann may not have had a satisfactory solution, but — unlike Prigogine, I think — he certainly saw the significance of the problem.

So Prigogine's criticism of Boltzmann seems misplaced. To my mind, it is also counterproductive. One of the hardest things in this field is to set off on the right foot — to decide just what needs explaining. Clear directions are hard to find, and *The End of Certainty* lives up to its name in an unintended way: it isn't to blame for current uncertainty, but seems likely to prolong it.

In Derek York's book, by contrast, lack of a clear direction turns out to be an advantage. *In Search of Lost Time* is an engaging ramble through diverse topics temporal, by a man who has been dating for 40 years and seems to have loved every minute of it. I wish he had provided more references for readers inclined to explore on their own, but otherwise this is a delightful little book. □

Huw Price is at the School of Philosophy, University of Sydney, NSW, Australia 2006.

Future present

Beyond Calculation: The Next Fifty Years of Computing

edited by Peter J. Denning and Robert M. Metcalfe

Copernicus: 1997. Pp. 313. \$24.30, £16.95

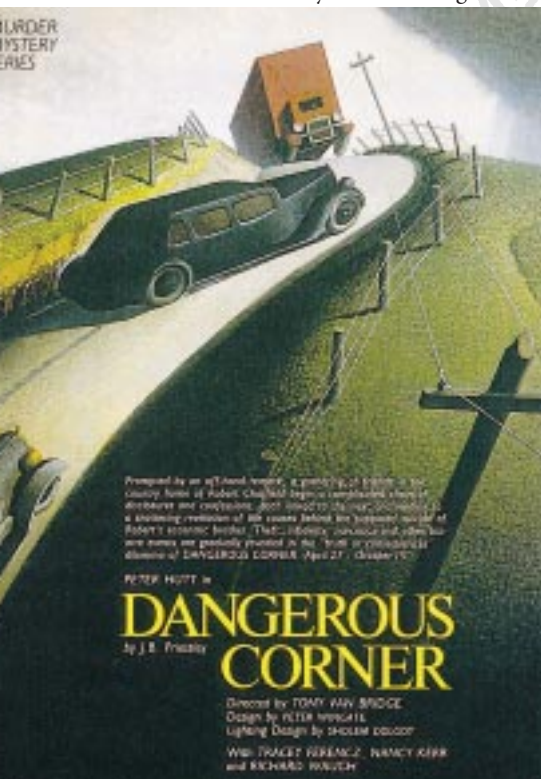
Martin Campbell-Kelly

How is this for an opener? "By 2047 almost all information will be in cyberspace — including a large percentage of knowledge and creative works. All information about physical objects, including humans, buildings, processes and organizations, will be online. This trend is both desirable and inevitable."

This statement comes from the opening essay of this book, written by Gordon Bell, perhaps the most influential computer architect of his generation, and James N. Gray, a database guru, both now senior researchers at Microsoft. Not many computer scientists or Web-surfers would dispute any of this, and it sets the tone for a compelling book. But don't be misled by the opening: what sets the book apart from the general run of technology-future books is the authority of its contributors and the tone of restraint that pervades it, when compared with the genre's usual ludicrous extrapolations.

For example, in the same essay Bell and Gray remind us that, despite the plummeting cost per bit and ever-increasing capacity of storage media, "paper is likely to be with us for ever... a lasting, irreplaceable graphical user interface. We know of no technology in 1997 to attack paper's broad use." What is more, they point out, acid-free paper has a longevity of more than 500 years, hundreds more than any known magnetic or optical storage medium.

Beyond Calculation is the edited proceedings of a conference organized in 1997 to celebrate the golden jubilee of the Association of Computing Machinery, the professional society of US information processing. The theme of the conference, as indicated by the book's title, is that computers have already gone far beyond numerical calculation, and have vastly further to go in the next half century. The conference organizers have done a fine job in assembling a star cast of leading computer technology researchers



J. B. Priestley's 1932 play *Dangerous Corner* illustrates, says Derek York, the key concept of chaos — sensitivity to initial conditions. In 1964, Priestley also published a book entitled *Man and Time*. (From a 1988 Shaw Festival playbill.)

and people on the fringe of computing. The 20 essays in this *fin de siècle* collection have been seamlessly edited, and only three or four of the papers are evident pot-boilers.

The book is divided into three parts, dealing respectively with technology, computing and human identity, and business and economic issues. Of the six chapters in the technology stream, one of the highlights is an essay by Mark Weiser and John Seely Brown, both senior researchers at the famous Xerox PARC in Palo Alto, California. In the 1970s PARC researchers came up with the graphical user interface that 15 years later, as Microsoft Windows, made Bill Gates's fortune. Now workers at PARC have moved on to what they call "calm technology" — when computers become invisible and disappear into the woodwork. They foresee "clocks that find out the correct time after a power failure, kids' toys that are ever-refreshed with new software and vocabularies, paint that cleans off dust and notifies you of intruders, and walls that selectively dampen sounds." Got all that, Bill?

The first of the six essays on computers and human identity is by the eminent social researcher Sherry Turkle, who reflects on what it means for kids to grow up in a culture of video games and simulation. "Fifty years ago," she reminds us, "a child's world was full of things that could be understood in simple, mechanical ways. A bicycle could be understood in terms of its pedals and gears and a wind-up car in terms of its clockwork springs." But lever the back off a Gameboy, and the child is confronted by an impenetrable chip. Thus children are led to psychological rather than mechanical explanations of their playthings; in effect, they perceive machines as having a kind of inner life — "more alive than a car, but less alive than a bacterium".

Another essayist on computers and human identity is Terry Winograd, once *enfant terrible* of artificial intelligence, but now with his feet planted firmly on the floor. His essay touches on the vexed issue of the boundaries of computer science. Today there are terrific tensions and frustrations in computing research because a researcher on interface design probably has more in common with social psychologists on the other side of the campus than with the computer theoreticians in the same building. Winograd supposes that there may eventually be a split between interface designers and computer scientists, rather like the complementary roles of architect and civil engineer.

The final set of essays deals with economic and business issues. I think everyone will take away from this book a particular vision that captures their imagination above the others. For me it was an essay on "Information Warfare" by Larry Druffel, former director of the Software Engineering Institute. We have all picked up the idea, from the popular

science press or voices in the air, that a new form of warfare is on the horizon. In this new blitzkrieg, taking out the enemy's information systems has the potential to do vastly more damage than the physical destruction caused by Second World War bombers. What would these weapons look like? One example Druffel cites is that of a 'rogue' program that floods the networks with bogus messages, impeding the free flow of information. The more advanced the economy, the greater the potential for devastation. Take out Albania's information systems and it probably wouldn't notice for three weeks. But do the same for the United States, 20 years down the road when it is totally dependent on the Internet, and the consequences would be awesome.

On a more upbeat note comes the delightful essay "There and Not There" by William Mitchell, the architect and cultural critic from the Massachusetts Institute of Technology, and Oliver Strimple, director of the Computer Museum in Boston. They ask why would anyone ever go out when everything they might want to see is on-line? In this articulate and optimistic essay, the authors persuasively argue that telepresence is no substitute for 'being there'. Television or a Web-browser simply cannot capture the full experience of going to a live pop concert, theatre performance, art gallery or soccer match. That indefinable, emotional something one gets at the real event they call 'aura'. We all knew it in our hearts, but they have given us a word for it. □

Martin Campbell-Kelly is in the Department of Computer Science, University of Warwick, Coventry CV4 7AL, UK.

Incomplete theorems

What is Mathematics, Really?

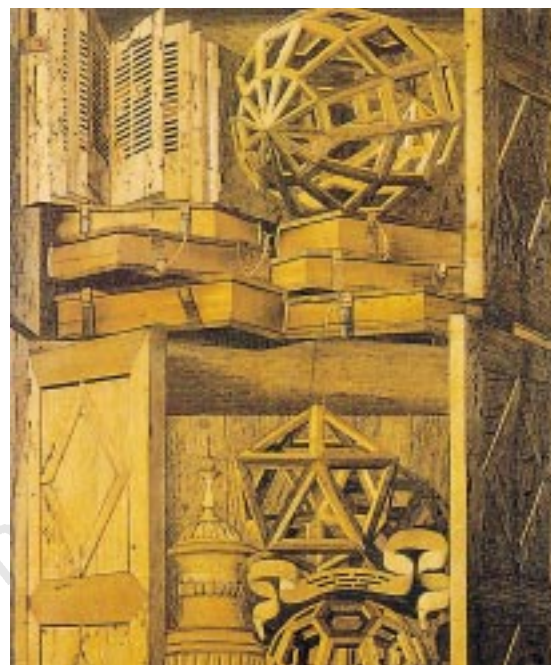
by Reuben Hersh

Oxford University Press/Cape: 1997. Pp. 343. \$35, £18.99

Ivar Ekeland

The title poses quite a question, and adding the word 'really' does not make it any easier. Empty words will not do: hard definitions are required. So what does 'really' really mean here?

It could mean: "What is modern mathematics, what makes it important in — or irrelevant to — our everyday life, what is it like to be an active mathematician?" There is much to say here. There has been an exponential growth in the number of scholars, books, journals, papers and theorems produced in the past century. Whole fields, such as probability theory, functional and numerical analysis, and noncommutative geometry, have been created. Older fields, such as algebraic



Facets of history: details of a wooden tarsia in the church of Santa Maria in Organo, Verona, Italy. From *Polyhedra* by Peter R. Cromwell (Cambridge University Press, £30, \$44.95).

geometry or number theory, have enjoyed tremendous progress, the solution of Fermat's last theorem being a case in point. The use of mathematics has become pervasive in engineering for filtering noisy signals and controlling processes, as well as in economics or finance, as shown by the success of the Black-Scholes model for pricing derivative securities.

It could also mean: "What is the inner meaning of mathematics?" Is there anything more to mathematics than its outward success story? Is it an activity to be pursued for its own sake, independently of material rewards? Painters or writers are supposed to be aiming for timeless achievement in art, not for recognition among their contemporaries, which may not even come during their lifetime. Is it the same with mathematicians? Are they artists in their own way? And, if so, what kind of art is theirs?

Strangely enough, this book goes neither way. It describes nothing of modern mathematics, except for a discussion about the foundations of mathematics and its connection with logic, which is far from the preoccupations of most working mathematicians. There is nothing at all to suggest the breadth and scope of mathematics today, or how much it has changed in the past century. It is always annoying to see the same tired, age-old examples crop up again and again: how to prove that the square root of two is irrational, how to use the tools of differentiation to exploit Galileo's law of falling bodies, and so on.

The book fails also to answer the question about the nature of mathematics. It dismisses it as "futile", leaving to forthcoming progress

in the neurosciences whatever kind of answer is possible. To quote the author: "How does mathematics come about, in a daily, down-to-earth sense? That question belongs to psychology, to the history of thought, and to other disciplines of empirical science. It can't be answered with philosophy."

So what does the author do? He presents his own philosophy of mathematics, which is now a fairly easy thing, considering all the questions he has decided such a philosophy should not answer. This philosophy he calls "humanism". The basic principle is: "A world of ideas exists, created by human beings, existing in their shared consciousness. These ideas have objective properties, in the same sense that material objects have objective properties. The construction of proof and counterexample is the method of discovering the properties of these ideas. This branch of knowledge is called mathematics."

I find it hard to understand what is meant by "shared consciousness" and to see how ideas could have "objective properties" in the "same sense" that material objects do. If I ventured such views in public, I would certainly try to clarify them, referring to such sciences as linguistics (after all, shared consciousness, whatever it is, must be mediated through language, and Chomsky has gone a long way towards a scientific analysis of the phenomenon) or to history (does a historical fact share the same objective properties as an experimental fact in physics?). Here, of course, such preoccupations are dismissed as "futile", so the author can take his "philosophy" for granted and proceed without further ado.

In the first part of the book, the author develops the ideas that "mathematics is human. It's part of and fits into human culture" and that "mathematics knowledge isn't infallible. Like science, mathematics can advance by making mistakes, correcting and recorrecting them." He contrasts these (very reasonable) ideas with a Platonist approach, which would take the stand that mathematics already 'exists' somewhere (written down in God's great book), so that theorems are discovered (unveiled), and not created, by humans. Mathematicians are then divided into humanists and Platonists, and the second part of the book is devoted to a review of prominent mathematicians through the ages, classified according to the stand they have taken on this question. At the end, their political opinions are reviewed as well, and, not surprisingly, even with the help of some elementary statistics, Platonists are found to be mostly conservative-righties whereas humanists are mostly democrat-lefties.

As a source of information about mathematics in general, the book is a failure. There remains scattered information, mostly in the form of quotations, about the foundations debate in mathematics. And the author does give unexpected entertainment at times. It is difficult to keep a straight face while reading

that Pythagoras, for instance, for whom no biographical data are available, except the facts that he lived in the sixth century BC and that none of his writings survive, was a 'rightie'. There is also some humour in seeing that the author, presumably wishing to spare the critics time and trouble, concludes his book with a "self-graded report card". The result: "Could be worse." Of course. □

Ivar Ekeland is at the Ceremade et Institut de Finance, Université Paris-Dauphine, 75775 Paris Cedex 15, France.

O brave new world that has Prozac in't!

Britain on the Couch: Why We're Unhappier Compared With 1950 Despite Being Richer – A Treatment for the Low-Serotonin Society

by Oliver James

Century: 1997. Pp. 402. £16.99

Chris McManus

Two subtitles maketh a reviewer's job easier. And should one miss the quasi-eighteenth-century prolixity of the title page, the opening lines also summarize the argument: "The premise of this book is that we are unhappier compared with 1950... that people who are unhappy tend to have low levels of serotonin and that levels thereof are largely caused by our social psychological environment."

Then follows a rambling mishmash of reviews of published research (always with hundreds of studies, thousands of subjects and unanimity of scientific interpretation), supplemented by case-histories from the author's experience, or his television documentaries, or secondary analysis of dysfunctionality in the British royal family. Only the latter, coupled with refreshing references to British rather than US television, justifies "Britain" in the title. As worldwide sales must be reduced, and the scientific evidence is international, I wondered about parallel editions for other countries — say, *Albania on the Couch*, detailing dysfunctionality in King Zog's descendants?

James's main argument is that we ("This book is about... people like us") are now more

depressed (as well as anxious, phobic, obsessional, anorectic, bulimic, gambling-addicted, violent, paranoid, alcoholic, drug-hooked and sexually promiscuous) than in 1950, the common denominator being reduced brain serotonin (present mainly as metaphor rather than molecule). Society has changed since 1950, becoming more competitive, less certain and less predictable. We know more about other people, and compare ourselves more, often with high-profile media-inflated role models, and find ourselves wanting. And a war between the sexes ("gender rancour" in James's quaint phrase), with less well-defined roles, increases divorce and disrupts childhood. This potent mixture makes naked apes, recently out of Africa, more depressed (and anxious, and so on). Is this plausible?

The book's argument depends critically on depression being more common now than half a century ago. Two other current controversies indicate the methodological problems. Is childhood asthma truly increasing, or is there just increased willingness to report symptoms or make diagnoses? Has intelligence really increased this century (scores on identical tests are certainly higher), or is there just increased impulsivity, test-wisness, guessing or visual literacy?

For psychiatry the problems are much greater. Standardized instruments such as the General Health Questionnaire now routinely use higher cut-offs for 'caseness', as people more willingly acknowledge problems. Historical studies of depression have not used equivalent criteria. James's brief appendix on "the scientific evidence" depends almost entirely on retrospective self-diagnoses in Klerman's controversial study of subjects of different ages. Even if depression has increased, proving causation is even more problematic; we see what we want to see, particularly when social and political factors are involved.

The final chapters give James's prescriptions for raising our low serotonin levels. Twenty million of the UK population would benefit from that contemporary soma Prozac (despite the claimed low libido, erectile failure or anorgasmia in 30–70 per cent of users). Psychotherapy and a more collectivist, communitarian "advanced capitalism" would also help.

Neither is diet neglected, with a serotonin-boosting recipe reminiscent of George Bernard Shaw's crankiness: "One approach is to consume only the juice extracted from pears, sweet beetroot... and carrots, from a juice extractor for a period of three days every month, consuming as many apples as are required if hungry in the interim." Surprisingly, James — unlike G. B. S. — doesn't tell us to stimulate the phagocytes. □

Chris McManus is at the Centre for Health Informatics and Multiprofessional Education, University College London Medical School, Whittington Hospital, London N19 5NF, UK.



Mutation of the mouse *klotho* gene leads to a syndrome resembling ageing

Makoto Kuro-o*, Yutaka Matsumura*†, Hiroki Aizawa*†, Hiroshi Kawaguchi‡, Tatsuo Suga†, Toshihiro Utsugi†, Yoshio Ohyama†, Masahiko Kurabayashi†, Tadashi Kaname§, Eisuke Kumell, Hitoshi Iwasaki||, Akihiro Iida¶, Takako Shiraki-Iida*¶, Satoshi Nishikawa#, Ryozyo Nagai†* & Yo-ichi Nabeshima*††

* Division of Molecular Genetics, National Institute of Neuroscience, NCNP, 4-1-1 Ogawahigashi, Kodaira, Tokyo 187, Japan

† The 2nd Department of Internal Medicine, University of Gunma School of Medicine, 3-39-22, Showa, Maebashi, Gunma 371, Japan

‡ Department of Orthopaedic Surgery, Faculty of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113, Japan

§ Institute of Molecular Embryology and Genetics, Kumamoto University School of Medicine, Kumakoto 862, Japan

|| Lead Optimization Research Laboratory, Tanabe Seiyaku Co. Ltd, 2-2-50 Kawagishi, Toda, Saitama 335, Japan

¶ Tokyo Research Laboratories, Kyowa Hakko Kogyo Co. Ltd, 3-6-6 Asahimachi, Machidashi, Tokyo 194, Japan

Pharmaceutical Research Laboratories, Kyowa Hakko Kogyo Co. Ltd, 118 Shimotagari, Nagaizumi, Sunto, Shizuoka 411, Japan

†† Core Research for Evolutional Science & Technology (CREST), JRDC and †† Institute for Molecular and Cellular Biology, Osaka University, 1-3 Yamada-oka, Suita, Osaka 565, Japan

A new gene, termed *klotho*, has been identified that is involved in the suppression of several ageing phenotypes. A defect in *klotho* gene expression in the mouse results in a syndrome that resembles human ageing, including a short lifespan, infertility, arteriosclerosis, skin atrophy, osteoporosis and emphysema. The gene encodes a membrane protein that shares sequence similarity with the β -glucosidase enzymes. The *klotho* gene product may function as part of a signalling pathway that regulates ageing *in vivo* and morbidity in age-related diseases.

Ageing can be defined as the age-related deterioration of physiological functions necessary for the survival and fertility of an organism¹. Common age-related diseases linked to these functions include arteriosclerosis, cancer, dementia and osteoporosis. To determine how these diseases come about is central to understanding human ageing.

One approach to investigating the mechanism of human ageing is to find genes that determine inherited premature-ageing syndromes and cause rapid development of multiple age-related diseases early in life. Candidate genes for some of these syndromes have been identified by positional cloning²⁻⁴ and found to encode a putative helicase that unwinds the DNA duplex and directly affects DNA replication and repair, explaining why patients acquire various gene mutations in their somatic cells⁵. From these findings, we concluded that the genetic mutations associated with premature ageing syndromes are mutation-causing mutations⁶. To determine whether accumulation of mutations can cause the development of multiple age-related diseases, it will be necessary to identify as many premature-ageing syndrome genes as possible and to find animal models of human ageing and identify their genetic features.

Here we describe a transgenic mouse with several age-related disorders caused by an insertional mutation of a transgene. Mice homozygous for the transgene show various phenotypes resembling those in patients with premature-ageing syndromes: arteriosclerosis, osteoporosis, age-related skin changes and ectopic calcifications, together with short lifespan and infertility. (We named this mutant *klotho*, for one of the Fates, the Greek goddess who spins the thread of life.) We have identified the gene linked to these ageing syndromes: it is not a helicase, but a new type of membrane protein. This indicates that the ageing phenotypes seen in the *klotho* mouse may be brought about by completely different mechanisms from those responsible for the premature-ageing syndromes.

Generation of the *klotho* mouse

Previously we reported salt-sensitive hypertension in transgenic

mice that overexpress the rabbit type-I sodium-proton exchanger⁷. In that study, 28 independent transgenic mice were produced by a standard microinjection method; three of the mice expressed the exogenous transgene but the other 25 strains did not. Each of the transgenic mice that did not express the transgene was independently mated in an attempt to obtain mice homozygous for the transgene-inserted allele and each was examined to determine whether any phenotypes caused by the insertional mutation had appeared. One of these mice, now termed *klotho*, exhibited interesting phenotypes that resembled human ageing; these phenotypes only appeared in mice homozygous for the transgene. Penetrance of all phenotypes was 100%. Because *kl/kl* mice develop normally up to at least 2 weeks of age in both macroscopic and histological appearance, the phenotypes seen in *kl/kl* mice cannot simply be a result of incomplete development.

Cumulative genotyping of heterozygous crosses revealed that the manner of transgene transmission was consistent with mendelian inheritance, demonstrating that intra-uterine or perinatal death of *kl/kl* mice had not occurred. The genetic background of the original *kl* mouse was a mixture of C57BL/6J and C3H/J. However, the phenotypes were the same for mice whose genetic background was replaced by BALB/c after repetitive backcrossing for more than 12 generations.

Ageing phenotypes in *klotho* mice

Growth retardation and short lifespan. Up to 3 to 4 weeks of age, *kl/kl* mice grow normally and are indistinguishable from their +/+ or *kl/+* littermates. At that time, *kl/kl* mice begin to show growth retardation (Figs 1a and 2b, c), gradually become inactive and marantic, and die prematurely at ~8–9 weeks of age. Their average lifespan is 60.7 days, with no *kl/kl* mouse ever surviving for longer than 100 days (Fig. 2a). The gross appearance of *kl/kl* mice was normal, except that they showed kyphosis (Fig. 1b). We were unable to specify the cause of death because each of the disorders described below is not in itself fatal.

Hypokinesia and gait disturbance. Open-field testing demonstrated that the spontaneous activity of *kl/kl* mice over 6 weeks of age was no more than 50% of that of control mice (Fig. 2d, e). Hind-paw footprint analysis revealed an abnormal walking pattern in *kl/kl* mice (Fig. 2f). Mean stride lengths corrected for base widths were significantly decreased compared with control mice, resembling the parkinsonian gait of the human aged⁸: altered function of the central nervous system may contribute to this hypokinesia and gait disturbance. Histological analysis of the central nervous system revealed that the number of Purkinje cells was decreased and that some of them had degenerated, but there were no other age-related changes such as brain atrophy, senile plaques or amyloid deposits in *kl/kl* mice.

Atrophy of genital organs and thymus. The external genital organs of *kl/kl* mice are atrophic and both sexes are incapable of mating (Fig. 1c, d). Macroscopic observation of *kl/kl* mice during dissection revealed atrophy of the testes, uterus and ovaries (Fig. 1e, f). The thymus was barely detectable in any *kl/kl* mice at 6–9 weeks old. This is not caused by a developmental defect: rather, it is a result of severe atrophy after normal development as the thymus was normal in size at earlier developmental stages. Atrophy of the thymus is widely observed during ageing of both humans and mice⁹.

Arteriosclerosis. In the aorta, we observed extensive medial calcification (Fig. 3a). In middle-sized muscular arteries, not only medial calcification but also intimal thickening were seen (Fig. 3b). Small arteries in the kidney were also extensively calcified (Fig. 3c). Arteriosclerosis first appears around 4 weeks after birth and progresses gradually with age. The vascular changes seen in *kl/kl* mice are very similar to those in humans that are found in arteriosclerosis of the Mönckeberg type, common in human ageing¹⁰.

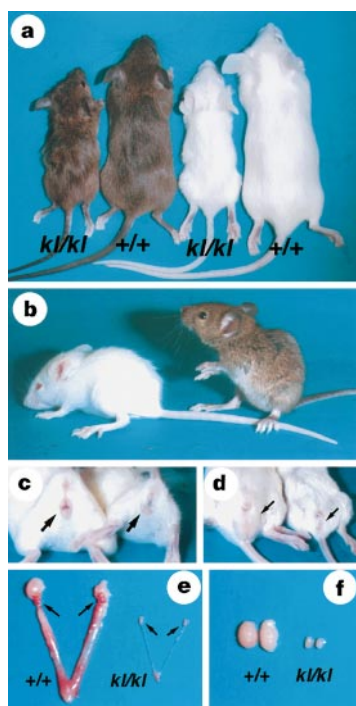


Figure 1 Macroscopic findings of *klotho* mice (8 weeks old). **a**, *kl/kl* mice with original (agouti) and with Balb/c background (albino) are shown with their *+/+* littermates. **b**, Close-up of *kl/kl* mice. **c**, Female external genital organs. The vaginal opening (arrows) is open in *+/+* mice (left), but closed in *kl/kl* mice (right). **d**, Male external genital organs. The scrotum (arrows) of *kl/kl* mice (right) is smaller than that of wild-type littermates (left). **e**, **f**, Female (**e**) and male (**f**) reproductive organs. Ovaries (arrows), uteri and testes are extremely atrophic in *kl/kl* mice.

Ectopic calcification. Ectopic calcification was evident in various organs of *kl/kl* mice as well as in arterial walls, including in the stomach (Fig. 7d), bronchial mucosa, alveolar cells, choroid plexuses, skin, testes and cardiac muscle. It also appears around 4 weeks after birth and progresses according to age. Overall, the distribution of ectopic calcification in *kl/kl* mice resembles that in natural human ageing¹¹.

Osteoporosis. In *kl/kl* mice, bone radiographs detected a generalized decrease in bone radiodensity, indicating the existence of osteopenia (Fig. 3d). This decrease in bone mineral density was particularly noticeable at the mid-portion of the long bones, where *kl/kl* mice had an approximately 20% lower density than control mice ($14.38 \pm 2.52 \text{ mg cm}^{-2}$ versus $19.19 \pm 0.74 \text{ mg cm}^{-2}$ in tibiae, and $21.40 \pm 1.79 \text{ mg cm}^{-2}$ versus $24.6 \pm 2.06 \text{ mg cm}^{-2}$ in femurs; mean $\pm$ s.d., $n = 5$), which is highly significant ($P < 0.01$). No sex differences were apparent. Histomorphometric analysis uncovered a decrease in the thickness of cortical bone. The number of osteoblasts and osteoclasts were decreased, suggesting a state of low turnover involving the formation and resorption of bone. All these findings bear a resemblance to senile osteoporosis in humans¹².

Skin atrophy. The hair of *kl/kl* mice was sparser than control mice. Histological examinations of the skin revealed a reduction in the number of hair follicles (Fig. 3e, f), another common feature of the aged¹³. There was also a reduction in dermal and epidermal thick-

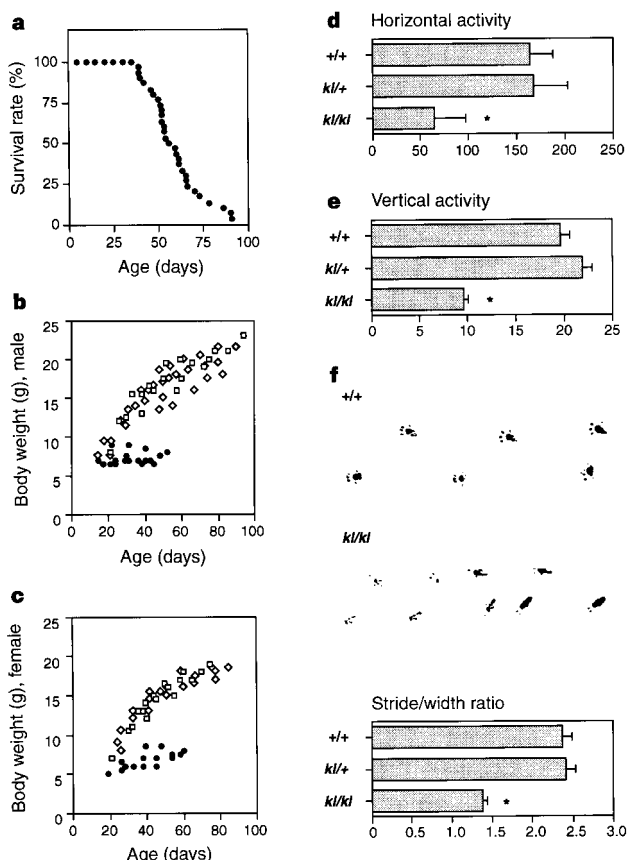


Figure 2 Short lifespan, growth retardation, and hypokinesia in *kl/kl* mice. **a**, Survival of *kl/kl* mice. The per cent survival *kl/kl* mice is plotted against age of the animals. **b**, **c**, Growth of male (**b**) and female (**c**) *klotho* mice. Body weight of *kl/kl* (circles), *kl/+* (diamonds), and *+/+* (squares) mice is plotted against age of the animals. **d**, **e**, Decrease in spontaneous horizontal (**d**) and vertical (**e**) activity detected by open-field test (see Methods). **f**, Hind-paw footprint test. Stride lengths corrected for the base widths (stride/width ratio) are decreased in *kl/kl* mice. Mean of the ten age- and sex-matched mice and s.d. (bars) were indicated. Asterisk, $P < 0.05$ versus *+/+* and *kl/+*.

ness. The subcutaneous fat was barely detectable, and, in general, the skin displayed overall atrophy. These findings are very similar to those found in senile atrophoderma in humans¹³.

Impaired maturation of gonadal cells. Both male and female *kl/kl* mice suffered impaired maturation of gonadal cells. In male *kl/kl* mice, the seminiferous tubules were extremely atrophic. Spermatoocytes failed to differentiate beyond the pachytene stage and therefore sperm did not mature (Fig. 3g, h). In the ovary, we found only immature follicles; there were no mature secondary follicles, Graafian follicles or corpus luteum (Fig. 3i, j), indicating that neither male nor female gametocytes could accomplish the first meiotic division.

Emphysema. Histology of the lungs of *kl/kl* mice was characterized by enlargement of the air spaces distal to the terminal bronchiole, accompanied by destruction of the normal alveolar architecture (Fig. 3k, l). These changes are identical to those found in emphysema in humans, the incidence of which increases with age¹⁴. The respiratory function of *kl/kl* mice was consistent with emphysema (data not shown), but the oxygen content in arterial blood was almost normal when breathing room air.

Abnormalities in the pituitary gland. Immunohistochemistry of the pituitary glands of *kl/kl* mice confirmed that growth hormone (GH)-producing cells were smaller than those of control mice. The luteinizing-hormone- and follicle-stimulating-hormone-producing cells also seemed slightly atrophic (data not shown). The number of secretory granules in GH-producing cells was significantly decreased in *kl/kl* mice (Fig. 3m, n). These findings strongly

suggest a defect in the signalling pathway for the appropriate production of pituitary hormones. Growth-hormone deficiency is known to contribute to the ageing process, resulting in growth retardation, reduced bone mass, reduced thymus weight, and reduced skin thickness¹⁵, all of which were also seen in *kl/kl* mice.

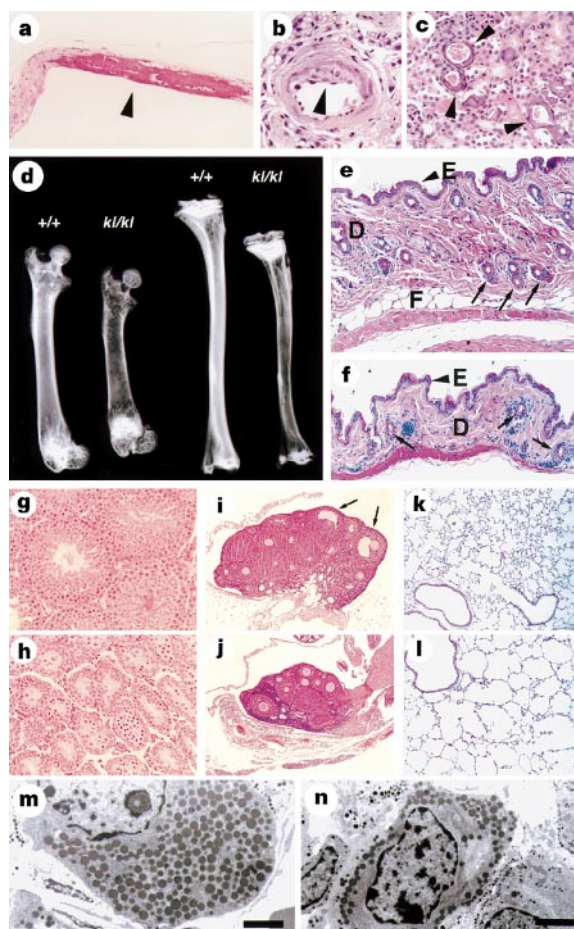


Figure 3 Histological analysis of *kl/kl* mice. **a-l**, Haematoxylin-eosin staining (not in **d**). **a**, Extensive medial calcification of the aorta (arrowhead, purple) in *kl/kl* mice. **b**, Intimal thickening in a muscular artery (arrowhead) in *kl/kl* mice. **c**, Calcification of the small arteries in kidney (arrowheads) in *kl/kl* mice. **d**, Bone radiographs of femurs (left two) and tibiae (right two). **e, f**, Skin of control (**e**) and *kl/kl* mice (**f**). Epidermal layer (E, arrowheads), dermal layer (D), subcutaneous fat layer (F), and hair follicles (arrows) are indicated. **g, h**, Testis of control (**g**) and *kl/kl* mice (**h**) at the same magnification. **i, j**, Ovary of control (**i**) and *kl/kl* mice (**j**). In control mice, there are mature secondary follicles (arrows). **k, l**, Lung of control (**k**) and *kl/kl* mice (**l**) observed in the same magnification. **m, n**, Electron micrograph of GH-producing cells in control (**m**) and *kl/kl* mice (**n**). Scale bar, 2 μ m.

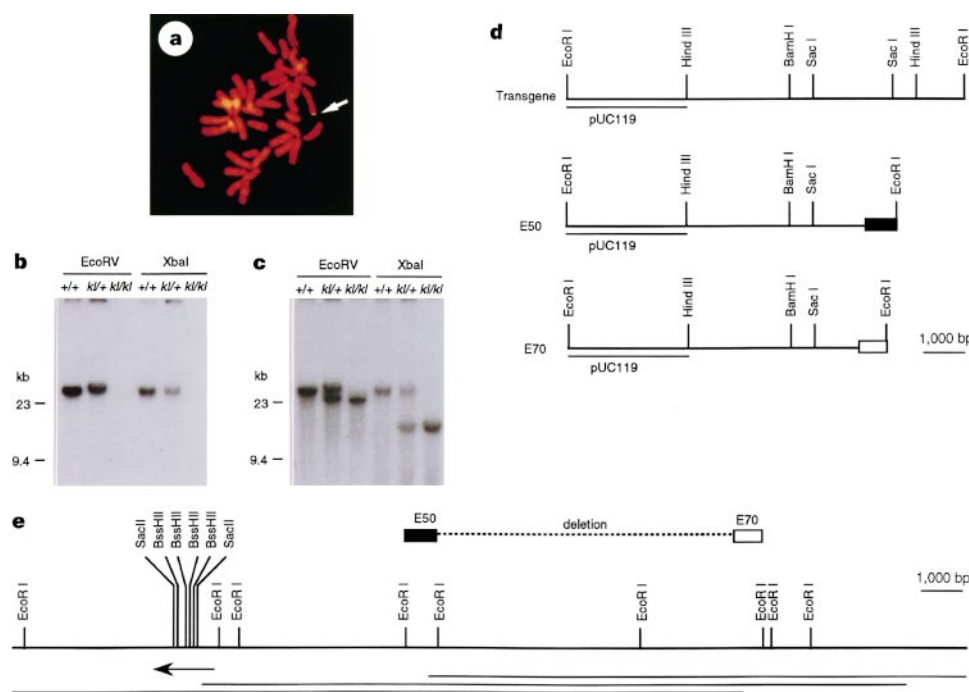


Figure 4 Identification of mouse *klotho* locus. **a**, Fluorescence *in situ* hybridization. The transgene insertion locus was determined as 5G3 (arrow). **b**, Southern blot analysis probed with the whole transgene. Genomic DNA from wild-type (+/+), heterozygous (*kl/+*) or homozygous (*kl/kl*) mice was digested with either *EcoRV* or *XbaI*. **c**, Southern blot analysis using the same filter as in **b**, using the E50-derived mouse genomic DNA fragment as a probe: the mutant (>50 kb) and wild-type alleles (25 kb by *EcoRV* digest and 12 kb by *XbaI* digestion) were detected. **d**, Restriction maps of the transgene and the rescued plasmids. DNA originating from the mouse genome is indicated (E50, black box; E70, white box). **e**, Physical map of the mouse *kl* locus. Horizontal bars at the bottom indicate the overlapping phage clones isolated. The positions of E50- and E70-derived mouse genomic DNA are indicated as black and white boxes, respectively. The direction of *kl* gene transcription is indicated by an arrow.

Given that the GH production in *kl/kl* mice is impaired, it may affect the ageing phenotype.

Serum and blood data. A slight increase in calcium ($9.47 \pm 0.30 \text{ mg dl}^{-1}$ versus $10.64 \pm 1.07 \text{ mg dl}^{-1}$, mean $\pm$ s.d. of $+/+$ and kl/kl mice, respectively; $n = 12$) and phosphorus ($8.54 \pm 1.34 \text{ mg dl}^{-1}$ versus $15.09 \pm 1.34 \text{ mg dl}^{-1}$) was observed, suggesting that these may contribute to ectopic calcification. Renal function seemed unaffected because creatinine levels were normal. *kl/kl* mice were hypoglycaemic ($231.5 \pm 22.6 \text{ mg dl}^{-1}$ versus $115.8 \pm 13.8 \text{ mg dl}^{-1}$), with decreased insulin in the pancreas, although it is not clear why. Peripheral blood analysis showed that the ratio of lymphocytes to leukocytes was decreased in *kl/kl* mice ($61.3 \pm 14.1\%$ versus $34.5 \pm 7.6\%$), consistent with atrophy of the thymus. Other data, including total protein, albumin, cholesterol and triglyceride levels, were normal in *kl/kl* mice. These results prove that the phenotypes of *kl/kl* mice are not derived

from such disease conditions as malnutrition, abnormal lipid metabolism, or chronic renal failure.

Cloning of the *klotho* gene

Approximately 20 copies of the transgene were integrated into the *kl/kl* mouse genome (data not shown). The chromosomal localization of the transgene insertion site was determined by the fluorescence *in situ* hybridization (FISH), using the whole transgene as a probe, which resulted in a single pair of symmetrical signals (Fig. 4a). In genomic Southern blot analysis using the whole transgene probe, only single bands were detected when genomic DNA was digested with restriction enzymes that do not cut the transgene (Fig. 4b). These results verify that multiple copies of the transgene were integrated in tandem at a single locus.

To generate the *klotho* mouse, the transgene was prepared by linearizing the transgene plasmid without removing the vector DNA (pUC119) (Fig. 4d). This allowed the convenient isolation of the DNA flanking the insertion locus by plasmid rescue (see Methods). Two kinds of plasmids (termed E50 and E70), each containing a short unique sequence of DNA that is not a part of the transgene, were recovered (Fig. 4d). Southern hybridization using these unique portions as probes indicated that they originated from the mouse genomic DNA flanking both ends of the transgene insertion site (Fig. 4c). A wild-type mouse genomic library was screened using the flanking DNA fragment as a probe. Three overlapping phage clones were isolated, covering about 25 kilobases (kb) of mouse genome (Fig. 4e). Both the E50 and E70 unique sequences were assigned to this region, separated by about 8 kb, indicating that the integration of the transgene produced a deletion of about 8 kb in the *kl* locus.

To identify exons in the *kl* locus, we determined the genomic sequence of the wild-type allele and analysed it with the exon-finding program GRAIL¹⁶. One of the GRAIL-predicted exons, located about 6 kb from the deletion site, contains multiple *Bss*HII (GCGCGC) and *Sac*II (CCGCGG) restriction sites, suggesting the existence of a CpG island (Fig. 4e). The 450-base-pair (bp) *Sac*II genomic fragment from this region was used as a probe for northern blot analysis and turned out to hybridize to an RNA fragment. The 5.2-kb transcript was detected predominantly in the kidney and faintly in the brain (Fig. 5a). Northern blot analysis of kidney and brain RNA from *kl/kl* mice showed that the *kl/kl* mouse is a null strain for this transcript (Fig. 5b). However, the more sensitive polymerase chain reaction with reverse transcription (RT-PCR) was able to detect its expression (Fig. 5c), indicating that the *kl* mutation was not a null but a severe hypomorph. We have determined the structure of the *kl* gene and found that it was in the 5' upstream region that the deletion occurred in the mutated allele (data not shown). Therefore, the *kl* gene may be slightly transcribed in *kl/kl* mutants.

A mouse kidney cDNA library was screened with the 450-bp *Sac*II fragment, enabling the full-length cDNA to be isolated. The protein predicted from the cDNA sequence is 1,014 amino acids long and contains a putative signal sequence at its N terminus¹⁷ and a single transmembrane domain near its C terminus, indicative of a new type-I membrane protein (Fig. 6a, b). We have confirmed by western blot analysis, indirect immunofluorescence, and fluorescent flow-cytometry that the recombinant KL protein is localized on cell surfaces when expressed in CHO cells (data not shown), which is consistent with the predicted KL protein structure.

The extracellular domain is composed of two internal repeats (termed mKL1 and mKL2), each about 450 amino acids long, which exhibit a weak similarity to each other (21% amino-acid identity). Each internal repeat shares homology with the β -glucosidases of both bacteria and plants and with a lactase-phlorizin hydrolase of mammals^{18,19}; the amino-acid identity is between 20 and 40%. It remains to be determined whether the KL protein has β -glucosidase activity.

To isolate a human homologue of the *kl* gene, a human kidney

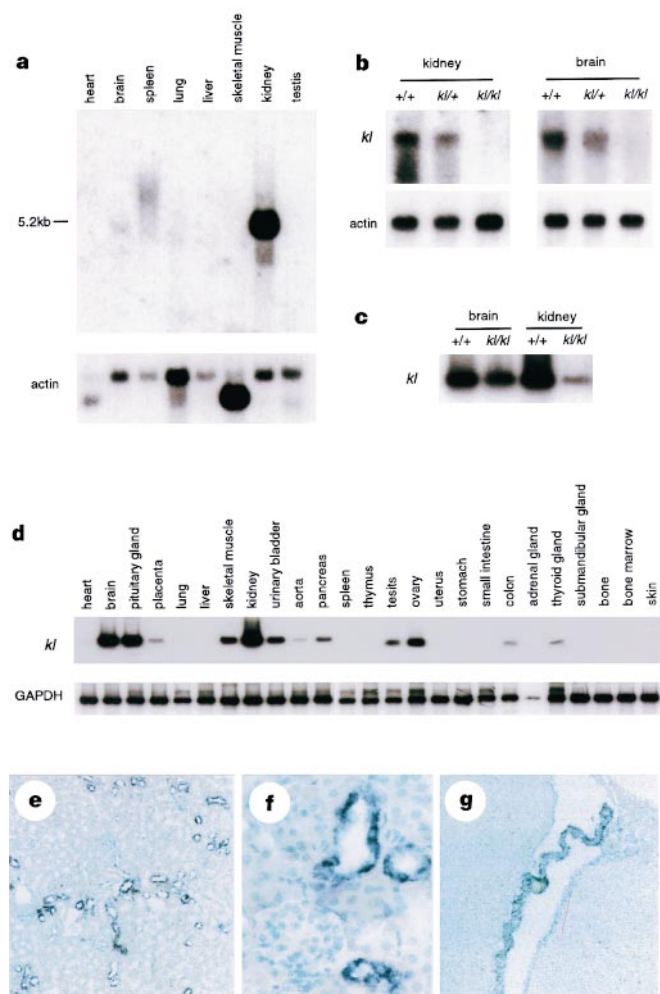


Figure 5 Expression of the mouse *klotho* gene. **a**, Northern blot analysis probed with the *Sac*II 450-bp genomic fragment from the *kl* locus (top panel) and human β -actin (bottom panel). Each lane contains 2 μ g poly(A)⁺ RNA (mouse MTN blot; Clontech). **b**, Northern blot analysis of poly(A)⁺ RNA from kidney (left) and brain (right) of wild-type (+/+), heterozygous (*kl*/+) and homozygous (*kl/kl*) mice. **c**, RT-PCR analysis of mouse *kl* mRNA expression. PCR products were transferred onto a nylon membrane after electrophoresis and detected by a ³²P-labelled *kl* cDNA fragment. **d**, Top, expression of the *kl* gene in various tissues detected by RT-PCR, as in **c**; bottom, expression of mouse glyceraldehyde 3-phosphate dehydrogenase (GAPDH) detected by Control Amplimer Set (Clontech). **e–g**, *In situ* hybridization analysis of mouse *kl* mRNA expression in kidney (**e**, lower magnification; **f**, higher magnification) and brain (**g**) of control mice. Sense probe gave no signal (data not shown).

from wild-type mice. They were first mated with *kl/+* mice to obtain transgenic mice carrying the heterozygous *kl* mutation in the F_1 generation; these mice were then backcrossed with *kl/+* mice and, in the F_2 generation, transgenic mice carrying the *kl/kl* homozygous mutation were obtained and their phenotypes analysed.

Out of 37 transgenic mice, two independent strains, numbers 46 and 48, apparently rescued all of the macroscopic phenotypes observed in *kl/kl* mice. The *kl/kl* mice bearing the transgene were almost indistinguishable from their wild-type littermates in appearance and growth (Fig. 7a), and both males and females were fertile. The thymus and genital organs were also restored to nearly normal weights. The serum levels of calcium, phosphorus and glucose were restored to almost normal values (data not shown). Histological analysis revealed that various pathological findings were dramatically ameliorated or completely normal in the rescued *kl/kl* mice. Arteriosclerosis was markedly improved (Fig. 7b, c). Ectopic calcification was considerably less than in *kl/kl* mice (Fig. 7d, e). In the tibia and femur, the thickness of cortical bone was normal and osteopenia was improved (Fig. 7f, g). The other phenotypes, including emphysema, atrophy of GH-producing cells, impairment of gonadal cell maturation, and skin changes, were also eliminated (data not shown). Thus, all disorders observed in *kl/kl* mice were improved by exogenous *kl* gene expression according to macroscopic, histological and blood analyses. These results confirm that *kl* is the gene responsible for the mouse *klotho* phenotypes.

Exogenous *kl* gene expression was never detected by RNase protection assay in transgenic lines that failed to rescue *kl/kl* mice. In strain 46, the transgene was expressed in tissues such as the brain, lung, liver, testis and ovary, as expected. On the other hand, in strain 48, transgene expression was detected only in brain and testis (data not shown). Expression of the exogenous gene was never detected in kidney, where endogenous expression is predominant. In addition, all the systemic phenotypes of *kl/kl* mice were improved even when exogenous *kl* gene expression was limited to the brain and testis. These results can be interpreted in two ways: first, expression in the kidney may not be essential to the function of the *kl* gene; second, the pleiotropic effects of *kl* gene product may be cell non-autonomous.

Discussion

We have established a novel mouse autosomal recessive mutant, *klotho*, that exhibits multiple phenotypes very similar to those observed during human ageing. Although some common phenotypes seen in natural human ageing, such as the development of tumours or cataracts, are not seen in *kl/kl* mice, the *klotho* mouse can be regarded as a model of human premature ageing syndromes because the phenotypes displayed by them fulfil many of the pathophysiological criteria for human ageing²³. No known laboratory mouse develops the syndrome observed in *kl/kl* mice, regardless of how long they live. Therefore, we should consider *kl/kl* mice as a model not for mouse ageing, but rather for human progeroid syndromes.

Senescence-accelerated mice (SAM) and their substrains have been developed for the study of human ageing and are known to exhibit ageing phenotypes, such as amyloidosis, osteoporosis and cataracts, among others²⁴. The *klotho* mouse differs from SAM in several respects: first, the multiple ageing phenotypes in *kl/kl* mice are autosomal recessive and uninfluenced by genetic background, whereas the conditions of inheritance in SAM are more complex and have not yet been clarified; second, the multiple ageing phenotypes identified in *kl/kl* mice all appear in these mice, whereas the different phenotypes associated with SAM are distributed among the various SAM substrains; and last, the ageing phenotypes in *kl/kl* mice manifest much earlier than in SAM. We assume that multiple gene mutations are implicated in causing the phenotypes observed in SAM.

When possible functions are considered for the KL protein, it

must be kept in mind that the *kl* gene is mainly expressed in specific cells and tissues. Some organs severely affected in *kl/kl* mice were not found to support *kl/kl* gene expression. In addition, the rescue experiment demonstrated that the exogenous *kl* gene expressed in limited organs could improve systemic ageing phenotypes in *kl/kl* mice. To explain these apparently non-cell-autonomous phenomena, it must be assumed that humoral factor(s) mediate the pleiotropic functions of the KL protein, at least in part. Our study indicates that ageing *in vitro* may be regulated through a humoral signalling pathway. Further investigation, including parabiosis experiments, will be needed to prove the existence of such humoral factor(s).

To our knowledge, the *klotho* mouse is the first laboratory animal model with multiple phenotypes resembling human ageing caused by a single gene mutation. Analysis of their pathophysiology will give clues not only to understanding the individual disease mechanisms but also the relationship between these mechanisms, essential to any discussion of human ageing. □

Methods

Histological examination. Ten male and 10 female *kl/kl* mice aged 6–9 weeks old, 4 male and 4 female *kl/kl* mice aged 4 weeks old, and 2 male and 4 female *kl/kl* mice aged 2 weeks old, were examined histologically together with age- and sex-matched *kl/+* or *+/+* mice. Organs were excised, fixed with 10% formaldehyde, embedded in paraffin, sectioned in 4- μ m slices and stained with haematoxylin–eosin and von Kossa staining. The organs examined were liver, kidney, heart, lung, spleen, trachea, tongue, submandibular gland, oesophagus, stomach, small intestine, large intestine, rectum, brain, cerebellum, eye, pituitary gland, adrenal gland, thyroid gland, parathyroid gland, genitalia, urinary bladder, femur, tibia, knee joint, ankle joint, thigh muscle, skin and aorta. The lungs were perfused through the trachea at a constant pressure of 20 cm H_2O with the fixative before fixing.

Behavioural analysis. Spontaneous locomotor activity was measured using a square arena (50 cm by 50 cm) with lattice lines at 5-cm intervals on its floor. A mouse was placed in the centre of the arena and horizontal activity was recorded by counting the total number of squares (5 cm by 5 cm) into which a mouse put its hind legs during 5 min. Vertical activity was recorded simultaneously by counting its rearing. The hind-paw footprint test has been described²⁵.

Fluorescence *in situ* hybridization (FISH). FISH was done as described previously²⁶. Probes (the entire transgene for mouse and the cDNA fragment containing the entire open reading frame for human) were biotin-labelled by nick translation.

Southern and northern blotting. Southern and northern blotting were done according to standard methods²⁷. For Southern blotting, genomic DNA prepared from liver was digested either by *Xba*I or *Eco*RV, which do not cut the transgene. After electrophoresis, DNA was transferred to a nylon membrane (Hybond-N+, Amersham). For northern blotting, poly(A)⁺ RNA was prepared from various organs of *+/+*, *kl/+* and *kl/kl* mice with oligo(dT) columns (Pharmacia), run on a denaturing formaldehyde/0.8% agarose gel, and blotted onto a Hybond N+ membrane according to the manufacturer's protocol. Probes were labelled with ³²P with random priming, using a commercially available kit (Amersham). A mouse multiple tissue northern blot was purchased from Clontech.

Plasmid rescue. Genomic DNA extracted from the liver of *kl/kl* mouse was completely digested with *Eco*RI and self-circularized with T4 DNA ligase (DNA Ligation System, TaKaRa) at the concentration of 10 μ g ml⁻¹. To rescue plasmids from the possibly methylated transgene, the ligation mixture was transformed into competent cells defective in certain methylation-dependent restriction systems (XL1Blue MRF', Stratagene) and plated on standard LB-ampicillin plates. All colonies were harvested and the rescued plasmids recovered by standard alkaline lysis.

Genomic and cDNA library screening. A mouse (129SVJ) genomic library, constructed on lambda FIX II (Stratagene), was used for cloning the *kl* locus. About 1 $\times$ 10⁶ plaques were screened with a ³²P-labelled *Eco*RI–*Pvu*II 620-bp fragment from E50. Complementary DNAs were prepared from poly(A)⁺ RNA of both an 8-week-old C57BL/6 male mouse kidney and human kidney

(Clontech). RNAs were primed with both oligo(dT)₁₂₋₁₈ and random hexamers simultaneously, then converted into cDNA using a commercially available kit (GIBCO BRL). cDNA libraries were constructed on lambda ZAP II (Stratagene) phage vector using Gigapack II Gold packaging extract (Stratagene). Approximately 9.0×10^5 primary plaques were obtained from both mouse and human kidney cDNA libraries.

RT-PCR. Poly(A)⁺ RNA (500 ng) from various organs was reverse-transcribed with random hexamer (TaKaRa) and 5% of the reaction mixture was amplified with LA-Taq DNA polymerase (TaKaRa) using a specific primer pair for mouse *kl* cDNA (5'-CCTGGTCGACCATTTTCAG-3' and 5'-AGCACAAAGTCGACAGACTTCTGGC-3'). Conditions for amplification were 30 cycles of 94°C for 30 s, 56°C for 30 s and 72°C for 90 s.

In situ hybridization. *In situ* hybridization was done using digoxigenin(DIG)-labelled antisense and sense riboprobes prepared by *in vitro* transcription of a mouse *kl* cDNA fragment (*Clal*-*Xba*I, 356 bp) with DIG-UTP and either T3 or T7 RNA polymerase according to the manufacturer's protocol (Boehringer Mannheim). The hybridization signal was detected as a blue precipitate using alkaline phosphatase-labelled anti-DIG antibody and a substrate containing nitro-blue-tetrazoliumchloride (NBT)/X-phosphate. Nuclei were counterstained with methyl green.

Generation of transgenic mice for rescue. A 4.2-kb *Not*I fragment of the mouse *klotho* cDNA clone (the 5' *Not*I site was derived from a polylinker of the λZAPII phage) containing the complete open reading frame was blunted and subcloned between the human elongation factor EF-1α promoter (including -580 bp upstream from exon 1, the first intron, and exon 2 truncated just 5' of the ATG start codon) and the SV40 small-T poly(A) signal cassette. This plasmid was used for pronuclear microinjection after linearization by *Not*I digestion.

Received 17 June; accepted 17 September 1997.

1. Maser, E. J. in *Handbook of Physiology, 11: Aging* (ed. Masoro, E. J.) 3-24 (Oxford Univ. Press, New York, 1995).
2. Yu, C.-E. *et al.* Positional cloning of the Werner's syndrome gene. *Science* **272**, 258-262 (1996).
3. Ellis, N. A. *et al.* The Bloom's syndrome gene product is homologous to RecQ helicases. *Cell* **83**, 655-666 (1995).
4. Weeda, G. *et al.* A presumed DNA helicase encoded by ERCC-3 is involved in the human repair disorders xeroderma pigmentosum and Cockayne's syndrome. *Cell* **62**, 777-791 (1990).
5. Salk, D., Au, K., Hoehn, H. & Martin, G. M. Cytogenetics of Werner's syndrome cultured skin fibroblasts: variegated translocation mosaicism. *Cytogenet. Cell Genet.* **30**, 92-107 (1981).
6. Ellis, N. A. Mutation-causing mutations. *Nature* **381**, 110-111 (1996).

7. Kuro-o, M. *et al.* Salt-sensitive hypertension in transgenic mice overexpressing Na⁺-proton exchanger. *Circ. Res.* **76**, 148-153 (1995).
8. Wolfson, L. & Katzman, R. in *Principles of Geriatric Neurology* (eds Katz, R. & Rowe, J. W.) 75-88 (Davis, Philadelphia, 1992).
9. Hausman, P. B. & Weksler, M. E. in *Handbook of the Biology of Aging* (eds Finch, C. E. & Schneider, E. L.) 414-432 (Van Nostrand Reinhold, New York, 1985).
10. Lansing, A. I. in *The Arterial Wall: Aging, Structure, and Chemistry* 136-160 (Williams & Wilkins, Baltimore, 1959).
11. Robbins, S. L., Angell, M. A. & Kumar, V. in *Basic Pathology* 3-27 (Igaku-shoin/Saunders, Tokyo, 1981).
12. Melton, L. J. & Riggs, B. L. in *Osteoporosis: Etiology, Diagnosis and Management* (eds Riggs, B. L. & Melton, L. J.) 155-179 (Raven, New York, 1988).
13. Chuttani, A. & Gilchrist, B. A. in *Handbook of Physiology, 11: Aging* (ed. Masoro, E. J.) 309-324 (Oxford Univ. Press, New York, 1995).
14. Sparrow, D. & Weiss, S. T. in *Handbook of Physiology, 11: Aging* (ed. Masoro, E. J.) 475-483 (Oxford Univ. Press, New York, 1995).
15. Corpas, E. S., Harman, M. & Blackman, M. R. Human growth hormone and human aging. *Endocr. Rev.* **14**, 20-39 (1993).
16. Ueberbacher, E. C. & Mural, R. J. Locating protein-coding regions in human DNA sequences by a multiple sensor-neural network approach. *Proc. Natl Acad. Sci. USA* **88**, 11261-11265 (1991).
17. Heijne, G. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* **14**, 4683-4690 (1986).
18. Mantei, N. *et al.* Complete primary structure of human and rabbit lactase-phlorizin hydrolase: implications for biosynthesis, membrane anchoring and evolution of the enzyme. *EMBO J.* **7**, 2705-2713 (1988).
19. Grabnitz, E., Seiss, M., Rucknagel, K. P. & Staudenbauer, W. L. Structure of the β-glucosidase gene *bg*/A of *Clostridium thermocellum*. *Eur. J. Biochem.* **200**, 301-309 (1991).
20. Uetsuki, T., Nakao, A., Nagata, S. & Kaziyo, Y. Isolation and characterization of the human chromosomal gene for polypeptide chain elongation factor-1α. *J. Biol. Chem.* **264**, 5791-5798 (1989).
21. Kim, D. W., Uetsuki, T., Kajiyo, Y., Yamaguchi, N. & Sugano, S. Use of the human elongation factor 1α promoter as a versatile and efficient expression system. *Gene* **91**, 217-223 (1990).
22. Hanaoka, K., Hayasaka, M., Uetsuki, T., Fujisawa-Sehara, A. & Nabeshima, Y. A stable cellular marker for the analysis of mouse chimeras: the bacterial chloramphenicol acetyltransferase gene driven by the human elongation factor 1α promoter. *Differentiation* **48**, 183-189 (1991).
23. Martin, G. M. Genetic syndromes in man with potential relevance to the pathobiology of aging. *Birth Defects* **14**, 5-39 (1978).
24. Takeda, T., Hosokawa, M. & Higuchi, K. Senescence-accelerated mouse (SAM): a novel murine model of accelerated senescence. *L. Am. Geriatr. Soc.* **39**, 911-919 (1991).
25. Barlow, C. *et al.* *Atm*-deficient mice: a paradigm of ataxia telangiectasia. *Cell* **86**, 159-171 (1996).
26. Kaname, T. *et al.* Mapping basigin (BSG), a member of the immunoglobulin superfamily, to 19p13.3. *Cytogenet. Cell Genet.* **64**, 195-197 (1993).
27. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning* (Cold Spring Harbor Laboratory Press, New York, 1989).

Acknowledgements. We thank E. Ozawa, T. Ishikawa, K. Hanaoka and I. Nonaka for earlier contributions to this work, T. Matsuzaki, S. Kameya and Y. Hiroi for maintaining mice, and H. Yamato and Y. Nagai for bone analyses.

Correspondence and requests for materials should be addressed to M.K. (e-mail: kuroo@ncnaxp.ncnp.go.jp) or Y.N. (e-mail: nabemr@imch.osaka-u.ac.jp). The nucleotide sequence data will appear in the DDBJ, EMBL and GenBank nucleotide sequence databases under the accession numbers AB005141 and AB005142 for mouse and human mRNA for *klotho*, respectively.

YOURS TO HAVE AND TO HOLD BUT NOT TO COPY

The publication you are reading is protected by copyright law. Photocopying copyright material without permission is no different from stealing a magazine from a newsagent, only it doesn't seem like theft.

If you take photocopies from books, magazines and periodicals at work your employer should be licensed with CLA. Make sure you are protected by a photocopying licence.



The Copyright Licensing Agency Limited
90 Tottenham Court Road, London W1P 0LP
Telephone: 0171 436 5931 Fax: 0171 436 3986

A subsurface flow of material from the Sun's equator to its poles

P. M. Giles*, T. L. Duvall Jr†, P. H. Scherrer‡ & R. S. Bogart‡

* Stanford University, Department of Applied Physics, Stanford, California 94305-4085, USA

† Laboratory for Astronomy and Solar Physics, NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

‡ W. W. Hansen Experimental Physics Laboratory, Stanford University, Stanford, California 94305-4085, USA

Gas on the Sun's surface has been observed^{1–4} to flow away from the equator towards both poles. If the same flow persists to great depths, it could play an important dynamical role in the eleven-year sunspot cycle, by carrying the magnetic remnants of the sunspots to high latitudes⁵. An even deeper counterflow, which would be required to maintain mass balance, could explain why new sunspots form at lower latitudes as the cycle progresses⁶. These deep flows would also redistribute angular momentum within the Sun, and therefore help to maintain the faster rotation of the equator relative to the poles⁷. Here we report the detection, using helioseismic tomography, of the longitude-averaged subsurface flow in the outer 4% of the Sun. We find that the subsurface flow is approximately constant in this depth range, and that the speed is similar to that seen on the surface. This demonstrates that the surface flow penetrates deeply, so that it is likely to be an important factor in solar dynamics.

An axisymmetric flow in meridional planes is known in astrophysics as a meridional circulation. This type of flow in stars has a long history, having been first proposed in 1925 by Eddington⁸. In regions of stars that lack convective motions, the importance of these flows for stellar evolution is that they can mix different layers of the stellar material and hence destroy composition gradients. Of course these flows have not been observed directly in distant stars, as they have small magnitudes and are difficult to isolate without some resolution of the stellar disk.

For the Sun, the flow speed is small compared with the random motions at the surface (1 km s^{-1}) and with the solar rotation (2 km s^{-1}) and so has been difficult to observe. From the surface measurements, there is rough consensus that the flow is poleward in both hemispheres, with a peak magnitude of $10\text{--}20 \text{ m s}^{-1}$ at mid-latitudes. In principle, the depth dependence of the meridional circulation could be measured with the standard techniques of global-mode helioseismology, because such a flow should affect the frequencies of the solar p-mode oscillations. However, several other effects can mimic the type of frequency shift expected, including the centrifugal effects of rotation and the wave speed variations associated with the sunspot cycle. There is no way to isolate the signature of a meridional circulation from these other effects⁹, so other techniques must be employed. Earlier measurements of subsurface motions¹⁰ hinted that global poleward flow might be present beneath the surface as well, but the observations were confined to a small area on the Sun and hence did not constitute a truly global measurement.

Our measurements are based on the principles of time–distance helioseismology^{11–15}, a tomographic technique in which the time for acoustic waves to traverse subsurface ray paths is measured by using the temporal cross-correlation between the data at two points separated by distance Δ . The difference in travel times between waves travelling in opposite directions on the same ray path is sensitive to a flow with a component along the ray path, as a wave will travel downstream in a shorter time than the corresponding

wave travelling upstream. By using pairs of points on the same meridian but with different latitude separations, we obtain information on the meridional circulation as a function of both latitude and depth. The resulting signature of a flow is a unique interpretation as far as we know, except possibly in the presence of the strong magnetic fields of sunspots^{16,17}, which do not contribute significantly to the present study. A diagram showing how the rays penetrate into the Sun is shown in Fig. 1.

The analysis procedure will be described in full detail elsewhere, but briefly, full-disk dopplergrams were analysed for the month of June 1996, a little more than one rotation period. We used data from the SOI/MDI instrument on board the SOHO satellite¹⁸. The images are sampled once per minute and there are 10^6 pixels in each image. The month is divided into 8-h intervals with separate calculations done for each interval. The images are remapped on to a coordinate system with equal increments in longitude and sine of latitude, the increments being determined by the pixel separation near the disk centre. The cross-correlations in time of all pairs of points on the same meridian and separated by a surface distance of $< 73 \times 10^6 \text{ m}$ are computed for meridians separated from the disk centre meridian by less than 50° . The cross-correlations are then averaged over longitude and over the different 8-h intervals covering the month. The travel times are measured by fitting a Gabor wavelet to the cross-correlation function for 20 min centred on the average travel time for that surface separation¹⁹. The resulting time measurements depend on the phase speed of the waves.

Time differences (southward minus northward) are computed for the counterpropagating waves. Points with similar distance and latitude were averaged with resultant latitude resolution of 5° and four separate ranges of Δ covering $12\text{--}24$, $24\text{--}36$, $36\text{--}49$ and $49\text{--}73 \times 10^6 \text{ m}$. Results were also computed for an average of time differences over the entire distance range $12\text{--}73 \times 10^6 \text{ m}$. The maximum depth to which the instrument is sensitive is $26 \times 10^6 \text{ m}$, $\sim 4\%$ of the solar radius and 12% of the thickness of the convection zone. We show in Fig. 2 the time difference plotted against latitude averaged over all the distance ranges. The poleward motion in each hemisphere is clearly seen. The latitudinal behaviour is similar to the earlier surface measurements. In addition, a small offset of the curve below zero can be seen, $\sim -0.4 \text{ s}$ at the equator. This offset could be caused by a flow crossing the solar equator and heading southward with speed 4.9 m s^{-1} , or it could be caused by a position angle error of our instrument of 8 arcmin. However, if there is a position angle error it should not lead to a depth-dependent flow. At the time of the launch of the SOHO satellite, the alignment between the star trackers and our instrument was known to be better than 1 arcmin, but since launch we have had no

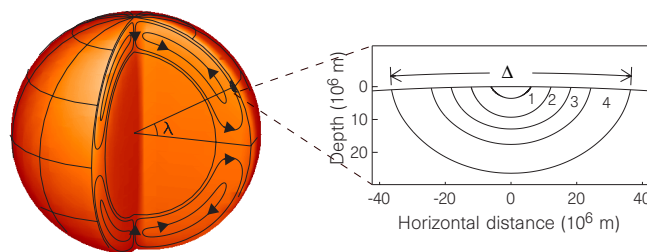


Figure 1 Diagram showing the geometry of the analysis. Examples of the ray paths at latitude λ are shown in the cutaway at the left (barely visible) and expanded on the right. The rays defining the boundaries of the four distance ranges are shown with numbers 1–4 defining the distance range. The deepest ray extends to 26×10^6 depth, $\sim 4\%$ of the solar radius ($R = 696 \times 10^6 \text{ m}$). The surface separation between the ray endpoints is defined to be the distance Δ . In this region of the Sun, the maximum depth of a ray is given approximately by Δ/π . Also shown in the cutaway is the lower boundary of the convection zone at $r = 0.7R$ and some streamlines of a simple model of a single-cell meridional circulation, with arrows indicating the flow direction.

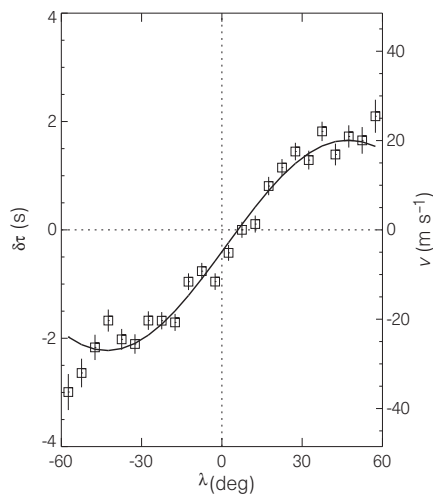


Figure 2 The average travel time difference (south minus north) for surface separation of pairs of points in the range $12\text{--}73 \times 10^6$ m. The individual points are shown (squares) and the 1σ errors (vertical lines). The solid curve is the best fit two-parameter model described in the text. From the modelling described in the text, we can put a velocity scale on the right axis, in which 12.1 m s^{-1} flow corresponds to a 1 s time difference.

reliable way of determining this angle. We have tried to use solar features to measure any misalignment, but if the meridional flow is different in the north and south hemispheres this will lead to an apparent angle error. We have therefore treated the signal as a flow, with the caveat that there could be a small offset caused by a position angle error.

We have fitted the travel time differences of Fig. 2 ($\Delta = 12\text{--}73 \times 10^6$ m) to a function of the form $\Delta\tau = a_1 \cos \lambda + a_2 \sin 2\lambda$, where λ is the latitude and a_1 and a_2 are estimated from the data by least squares. The first term is maximum at the equator, and attempts to measure the offset. This is approximately the form that we would expect if the offset were due to an orientation error. The second term fits the main part of the meridional flow; it is zero at the equator and both poles and peaks at 45° latitude. For $\Delta = 12\text{--}73 \times 10^6$ m, we get $a_1 = -0.41 \pm 0.04$ s and $a_2 = 2.02 \pm 0.06$ s. Fits were also attempted with a $\cos 4\lambda$ term as suggested by Durney²⁰, but its coefficient was not significant and has magnitude $a_4 < 0.21$ s (3σ). The best fit model is plotted with the data in Fig. 2. We have also measured these coefficients for the individual distance ranges described above. By studying how a_i varies for the different distances, we can learn about the depth dependence of the flow.

To do this requires some modelling of how the travel time difference is related to the velocity at depth. Our initial attempt is a fairly simple model, although we believe that it captures the essence of the problem. All of our modelling uses some very basic assumptions: that the waves are refracted by the changing sound speed and follow ray paths. A solar model is used to give the run of sound speed with depth²¹. In our simple ray model, the time difference between the oppositely propagating waves is given by $\Delta\tau = 2 \int \mathbf{v} \cdot \hat{n} ds/c^2$, where $\mathbf{v}$ is the flow velocity, which we assumed to be horizontal, $\hat{n}$ is the unit vector tangent to the ray, c is the sound speed, ds is the increment of length along the ray path, and the integral is along the ray path. Bogdan²² has shown that the response in depth is not as narrow as a ray, but is larger because of the finite wavelength of the acoustic waves. We have tried to take this into account by taking ray bundles (Fig. 1) that are already broad, and in fact the broadening expected is about the width of our bundles. We feel that our main conclusions are not affected appreciably by the small additional broadening expected.

We have compared two models with the data for the four

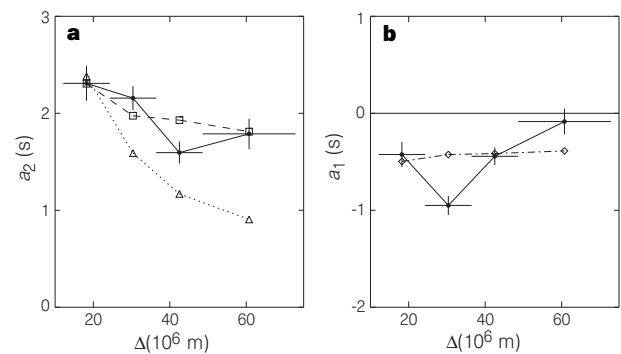


Figure 3 The fit coefficients a_2 (a) and a_1 (b) plotted against the distance Δ . The individual measurements for the four distance ranges are shown (solid circles connected by solid lines) with computed 1σ error bars and horizontal lines indicating the width of the distance range in both a and b. Also shown in a are the calculated coefficients for two different models of the flow velocity, both of which vary with latitude according to $v = v_0 \sin 2\lambda$. In one model (squares connected by dashed line), the velocity is constant with depth and the magnitude $v_0 = 23.5 \text{ m s}^{-1}$. In the second (triangles connected by dotted line), the flow is constant to a depth of 9.3×10^6 m, corresponding to our most shallow bin, and zero in the deeper layers. In this case the magnitude has been adjusted to $v_0 = 24.3 \text{ m s}^{-1}$. The observed points fit much better with the constant velocity model, suggesting that the flow extends throughout the observed depth range. In b a model is shown in which the apparent southward offset is due to an image orientation error of 8 arcmin (diamonds connected by dot-dashed line). There seems to be excess variation about this model, consistent with real solar flows. For those interested in further analysis, the four values of a_2 with errors are 2.31 ± 0.18 , 2.16 ± 0.12 , 1.59 ± 0.11 and 1.79 ± 0.16 s.

individual distance ranges. The first model has a constant velocity with depth. The second model has the motion concentrated in the top layer, corresponding to our most shallow range. Curves for these two models are plotted with the data for the four distances in Fig. 3. Clearly the uniform-velocity model is a much better fit than the surface-concentrated flow, indicating that the meridional circulation persists to the maximum depth of our analysis. Our best estimate is that the meridional flow is $23.5 \pm 0.6 \cos 2\lambda \text{ m s}^{-1}$, constant from the surface to 26×10^6 depth. We note that over this depth range the mass density ρ changes by several orders of magnitude, and so our results contradict some suggestions that the product ρv would be approximately constant.

Most, if not all, of the meridional flow should be in the outer 30% of the Sun. One goal of future work will be to increase sensitivity to be able to map this entire region, which we hope to do by also using waves propagating obliquely to the meridians. A second goal will be to observe the changes in this flow as the Sun progresses from the minimum of the sunspot cycle to a period of greater activity. Earlier measurements have indicated that the surface flow might be slower at sunspot maximum⁴, and in fact a slower flow is required by some models of the magnetic field⁵. Ultimately, we hope that with this technique we can gain a better understanding of the origin and dynamics of the Sun's magnetic cycle. $\square$

Received 4 August; accepted 25 September 1997.

1. Duvall, T. L. Jr Large-scale solar velocity fields. *Solar Phys.* **63**, 3–15 (1979).
2. LaBonte, B. J. & Howard, R. F. Solar rotation measurements at Mount Wilson. *Solar Phys.* **80**, 361–372 (1982).
3. Hathaway, D. H. et al. GONG observations of solar surface flows. *Science* **272**, 1306–1309 (1996).
4. Komm, R. W., Howard, R. F. & Harvey, J. W. Meridional flow of small photospheric magnetic features. *Solar Phys.* **147**, 207–223 (1993).
5. Wang, Y.-M., Nash, A. G. & Sheeley, N. R. Magnetic flux transport on the Sun. *Science* **245**, 712–718 (1989).
6. Dikpati, M. & Choudhuri, A. R. On the large-scale diffuse magnetic field of the Sun. *Solar Phys.* **161**, 9–27 (1995).
7. Gilman, P. A. & Miller, J. Nonlinear convection of a compressible fluid in a rotating spherical shell. *Astrophys. J.* **61**, 585–608 (1986).
8. Eddington, A. S. Circulating currents in rotating stars. *Observatory* **48**, 73–75 (1925).

9. Zweibel, E. G. & Gough, D. O. Is there a seismic signature of the Sun's magnetic field? in *Proc. Fourth SOHO Workshop Helioseismology* 73–76 (ESA SP-376, 1995).
10. Patrón, J., Hill, E., Rhodes, E. J. Jr, Korzenik, S. G. & Cacciani, A. Velocity fields within the solar convection zone: evidence from oscillation ring diagram analysis of Mount Wilson dopplergrams. *Astrophys. J.* **455**, 746–757 (1995).
11. Duvall, T. L. Jr, Jefferies, S. M., Harvey, J. W. & Pomerantz, M. A. Time–distance helioseismology. *Nature* **362**, 430–432 (1993).
12. Duvall, T. L. Jr, D'Silva, S., Jefferies, S. M. & Harvey, J. W. Downflow under sunspots detected by helioseismic tomography. *Nature* **379**, 235–237 (1996).
13. Kosovichev, A. G. Tomographic imaging of the Sun's interior. *Astrophys. J.* **461**, L55–L57 (1996).
14. D'Silva, S. Theoretical foundations of time–distance helioseismology. *Astrophys. J.* **469**, 964–975 (1996).
15. Duvall, T. L. Jr *et al.* Time–distance helioseismology with the MDI instrument: initial results. *Solar Phys.* **170**, 63–73 (1997).
16. Woodard, M. F. Implications of localized, acoustic absorption for heliotomographic analysis of sunspots. *Astrophys. J.* (in the press).
17. Braun, D. C. Time–distance sunspot seismology with GONG data. *Astrophys. J.* **487** (in the press).
18. Scherrer, P. H. *et al.* The solar oscillations investigation—Michelson doppler imager. *Solar Phys.* **162**, 129–188 (1995).
19. Kosovichev, A. G. & Duvall, T. L. Jr Acoustic tomography of solar convective flows and structures. in *Solar Convection and Oscillations and their Relationship* (eds Christensen-Dalsgaard, J. & Pijpers, F.) (Kluwer, Dordrecht, in the press).
20. Durney, B. D. On the solar differential rotation: meridional motions associated with a slowly varying angular velocity. *Astrophys. J.* **407**, 367–379 (1993).
21. Christensen-Dalsgaard, J. *et al.* The current state of solar modeling. *Science* **272**, 1286–1291 (1996).
22. Bogdan, T. J. A comment on the relationship between the modal and time–distance formulations of local helioseismology. *Astrophys. J.* **477**, 475–484 (1997).

Acknowledgements. T.D. thanks Phil Scherrer and the SOI group at Stanford for their hospitality during this work and for the use of their computing facilities, supported by a contract from NASA. This work was supported in part by the Solar Physics Branch of the Space Physics Division of NASA. SOHO is a mission of international cooperation between ESA and NASA.

Correspondence and requests for materials should be addressed to P.M.G. (e-mail: pgiles@solar.stanford.edu).

Metamorphosis of a quantum wire into quantum dots

Joel Hasen, Loren N. Pfeiffer, Aron Pinczuk, Song He, Ken W. West & Brian S. Dennis

Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974, USA

Bound states of electron–hole pairs (excitons) in semiconductors possess desirable properties—such as an enhanced oscillator strength for radiative recombination—that hold promise for the next generation of optical devices. However, at typical device operating conditions (room temperature and moderate charge densities), excitons dissociate to form an electron–hole plasma. Dissociation may be prevented by confining excitons to lower dimensions, where their binding energy is expected to increase significantly¹. But such confinement may in turn influence the dynamical properties of the excitons. Here we report spatially resolved photoluminescence images of excitons confined to an isolated gallium arsenide quantum wire. As the temperature of the structure is lowered, we observe a striking transition from broad and fairly continuous photoluminescence to an intense set of emission peaks which are both energetically sharp and spatially localized. Such behaviour indicates that, at sufficiently low temperatures, the quantum wire acts like a sparse set of quantum dots. Furthermore, at the site of an isolated quantum dot, we observe an unusual decrease in the relaxation rate of excitons, such that they radiate (via recombination) from higher energy states before relaxing to their ground state. We argue that this is the manifestation of an exciton relaxation ‘bottleneck’, the existence of which could pose problems for the development of optical devices based on quantum dots.

The quantum wires studied in this work are formed at the T-intersection² of 70-Å and 66-Å quantum wells (QW). In a QW, the wavefunction of an exciton is squeezed by the barrier potential and raised in energy. However, at the intersection of two QWs, the wavefunction can expand in the larger region, thereby lowering its energy, as shown in Fig. 1. Consequently, the excitons are confined to this intersection. Along the intersection, in an ideal structure they

are free to move, thereby forming a one-dimensional wire. These unique structures were grown using a two-growth molecular beam epitaxy (MBE) technique called cleaved-edge-overgrowth, described elsewhere³. A schematic cross-section of the sample is shown in Fig. 2a. Because the structure was grown using MBE, the cross-section of the quantum wire is expected to be uniform to within a few monolayers.

Previous studies on such structures have emphasized the lasing properties^{4,5} or the exciton energy shift^{6–9}. As a consequence, all previous photoluminescence (PL) measurements have been limited to samples containing a parallel array of closely spaced wires. In this work, we used high spatial and spectral resolution PL at low temperatures to image a single isolated quantum wire. This allows us to clearly match each PL peak with its corresponding location in the structure, and that has revealed striking new physics.

To maximize the spatial resolution of our measurements, we designed and built a low-temperature diffraction-limited confocal microscope. The samples were photo-excited with collimated laser light focused to the diffraction-limited spot of 0.8 µm full-width at half-maximum. Only the PL limited to the laser spot region was collected by using an optical fibre as an aperture in the image plane. Consequently, the PL from any excitons that diffused outside the laser spot region was blocked, as this light would have reduced the spatial resolution. Figure 2b shows the PL spectrum with the laser spot of 600 nW at 1,609 meV spatially coincident with the quantum wire at 20 K. The three broad peaks centred at 1,587, 1,576 and 1,565 meV correspond to excitons trapped in the stem QW, arm QW and quantum wire, respectively. The stem QW has the highest intensity because of its large depth of field in this collection geometry. To study the spatial properties such as wire cross-section and continuity, scans were made by moving the sample using a high-resolution x–y translation stage with a step size of 0.1 µm.

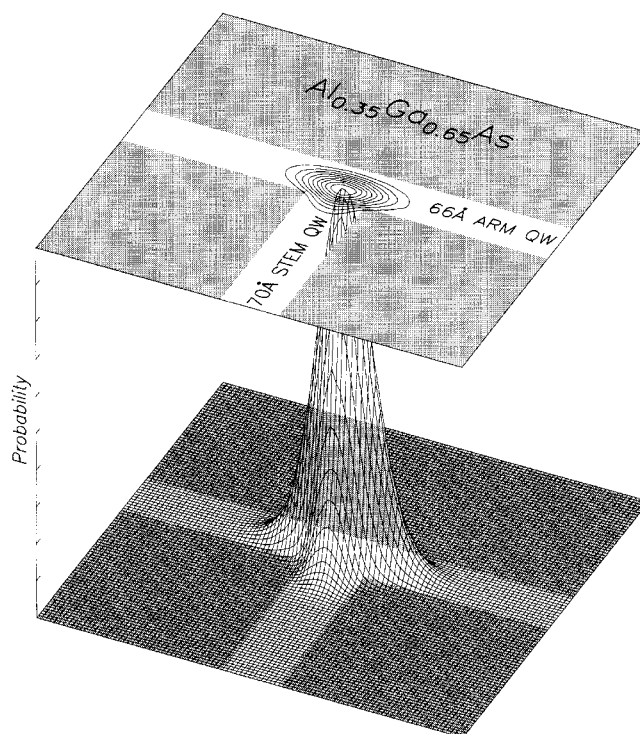


Figure 1 Schematic cross-section of the GaAs quantum wire, showing the centre-of-mass wavefunction of an exciton trapped at the T-intersection of the stem and arm quantum wells. The exciton is free to move along the intersection axis, thereby forming a one-dimensional wire. The wavefunction was obtained by solving the Schrödinger equation of an electron–hole pair using the fixed-node Monte Carlo method.

Figure 2c shows the PL obtained by scanning across the entire 10- μm -thick sample. The three PL peaks at position 2.0 μm identify the precise location of the quantum wire and correspond to the spectrum shown in Fig. 1b. Away from the quantum wire, the PL from the arm QW is visible across the entire sample as expected. Notice, however, that near position 4.5 μm the arm QW intensity is reduced. The low-energy 200- $\text{\AA}$ reference QW at 1,524 meV captures most of the excitons in that region and, consequently, depopulates the arm QW. The spot size is thought to be the resolution limit in far-field optics, but we argue that in the diffraction limit this is not the case. The focused laser spot has a gaussian intensity profile. Moreover, the PL from photo-excited excitons has a similar profile, an effect that is enhanced by the optical fibre which preferentially collects only the PL with a gaussian profile. As a result, the actual effective spatial resolution in our apparatus is enhanced to about 0.2 μm , as can be inferred from our ability to spatially locate PL peaks shown below.

A line scan of the PL along the quantum wire is shown in Fig. 3. In this case, there are three bands of PL corresponding to the stem QW, arm QW and quantum wire. The energy of the calculated exciton state of the quantum wire shown in Fig. 1 is 1,564.5 meV, in agreement with the measured value in Fig. 3. Clearly, over the 100- μm line scan, these three peaks are each fairly continuous on a macroscopic length scale. On a microscopic length scale, however, we observe local inhomogeneities in the PL of the quantum wire. The broad wire peak is accompanied by spectrally sharp peaks of higher intensity. These local features can be more clearly seen in the wire PL in Fig. 4 which is a series of line scans along the same 30- μm -long segment of the quantum wire at various temperatures. At 4 K, the quantum wire PL is transformed into an intense set of spectrally sharp, spatially localized peaks. The spectral widths of the peaks shown here are limited by the 200- μeV instrumental resolution, but other measurements conducted on a higher resolution spectrometer indicate that these peaks range from 80 to 150 μeV wide. Assuming that these peaks are homogeneously broadened, their lifetime could exceed 50 ps. The spatial widths of the peaks match the estimated gaussian intensity profile and are therefore indistinguishable from point sources located along the quantum

Figure 2 a, A schematic cross-section of the quantum wire (QWR) and related quantum wells. Laser light is focused to a spot on the arm QW. The sample was grown using a two-growth molecular beam epitaxy (MBE) technique called cleaved-edge-overgrowth. In the first MBE growth, the following sequence of layers was grown on a (100)-oriented GaAs wafer: (1) 0.5 μm GaAs buffer, (2) 1.1 μm $\text{Al}_{0.35}\text{Ga}_{0.65}\text{As}$ barrier, (3) 70 $\text{\AA}$ GaAs **stem QW**, (4) 2.2 μm $\text{Al}_{0.35}\text{Ga}_{0.65}\text{As}$ barrier, (5) 200 $\text{\AA}$ **reference QW**, and (6) 6 μm $\text{Al}_{0.35}\text{Ga}_{0.65}\text{As}$ barrier. The sample was removed from the chamber and returned to the chamber mounted in the (110) edge orientation. Upon reaching the desired growth conditions, the sample was cleaved leaving behind a virgin (110) surface. In the second MBE growth, a 66- $\text{\AA}$ GaAs **arm QW** was deposited followed by a 1- μm $\text{Al}_{0.35}\text{Ga}_{0.65}\text{As}$ barrier. **b**, The PL of the quantum wire region at position 2.0 μm . There are three peaks corresponding to the quantum wire, arm QW, and stem QW, respectively (bottom to top). **c**, Line scan of PL across the sample at 20 K corresponding to the structure in **a**.

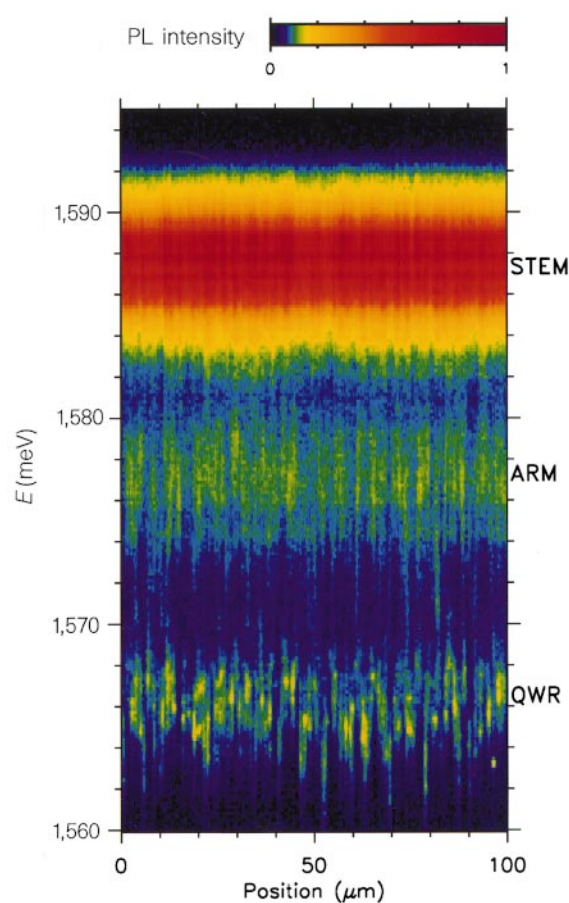
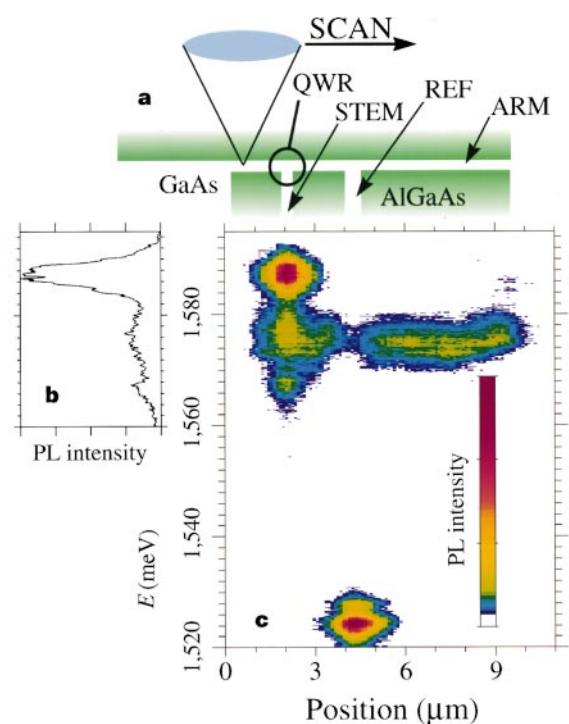


Figure 3 Line scan of PL along the quantum wire at 20 K. The peaks at 1,588, 1,576 and 1,565 meV correspond to the stem QW, arm QW and quantum wire, respectively. On a macroscopic length scale, these peaks are continuous as expected for a quantum wire. On a microscopic length scale, however, there are local inhomogeneities in the PL of the quantum wire.



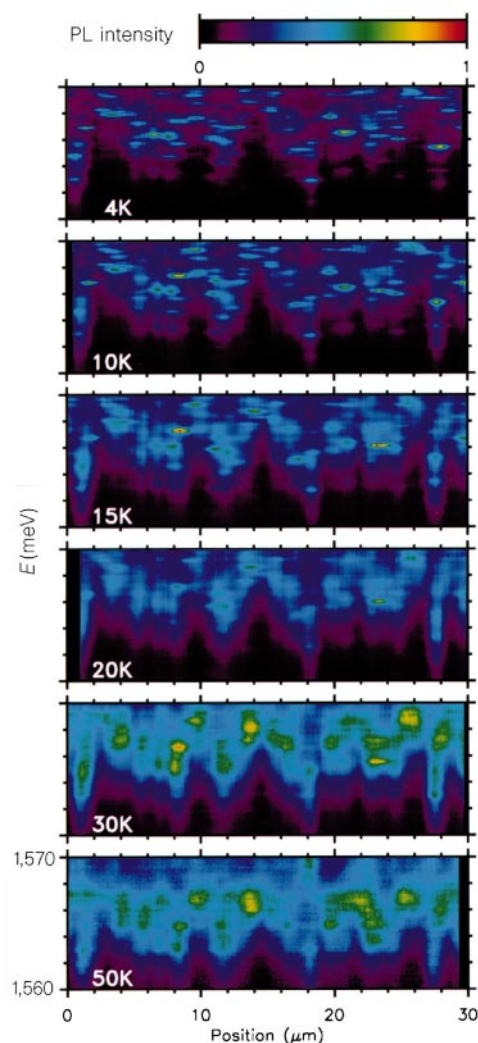
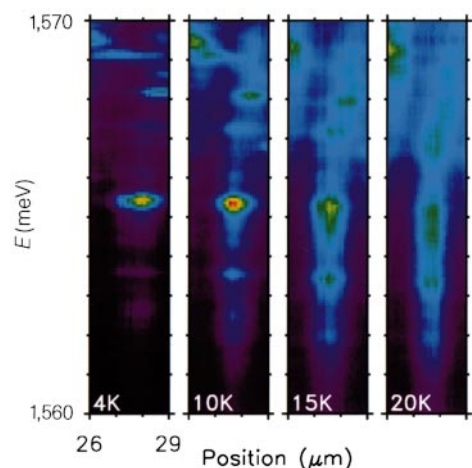


Figure 4 Line scan of PL along a 30- μm length of the quantum wire at several temperatures. At low temperatures, the PL is transformed to about 70 spectrally sharp, spatially localized peaks. These peaks are a signature of excitons localized in quantum dot states distributed along the wire, probably created by monolayer thickness fluctuations. At higher temperatures, the sharp peaks diminish in intensity and the PL is mostly a broad peak characteristic of quantum wire states with residual longer-period thickness fluctuations.



wire. It thus appears that the quantum wire is behaving as a set of quantum dots. We believe these states originate from the expected monolayer thickness fluctuations of the two intersecting QWs that make up the quantum wire. During the MBE growth of GaAs, large surface mobility allows the GaAs to form islands, 100 to 1,000 Å in diameter, of monolayer step height. These islands result in locally wide regions in the QW. The excitons are effectively confined to these wide regions, as the exciton wavefunction can lower its energy by expanding. At the T-intersection of the two quantum wells, the wide regions form quantum dots. The energy level shift in a quantum wire resulting from a single monolayer thickness fluctuation of either QW was estimated to be about 1.2 meV by numerically solving the two-dimensional Schrödinger equation. This agrees with the energy level spacing between quantum dot states observed in Fig. 4 at low temperatures. An actual count of the spectrally sharp peaks in Fig. 4 (4 K) gives a linear density of $2.3 \mu\text{m}^{-1}$. However, the quantum dot density is probably smaller because some of the peaks may be excited states of the same dot, as will be discussed below. As we raise the temperature, the sharp peaks diminish in intensity. By 50 K, the PL spectrum dramatically changes to a broad peak (5 meV wide). At these higher temperatures, it seems that the excitons have enough thermal energy to escape from the weak quantum dot confinement and migrate along the quantum wire. The remaining inhomogeneities are probably due to long-scale thickness fluctuations. The boundary between these two behaviours occurs at about 25 K which corresponds to an activation energy barrier of 2.1 meV. By our estimation, this barrier corresponds to a total fluctuation of two monolayers apportioned between the stem and arm QWs.

Recent physics literature contains several reports of spectrally sharp peaks from quantum dots in narrow QWs at low temperatures^{10–13}. Those QW dots also appear to originate from local monolayer thickness fluctuations. As those measurements were in much narrower QWs (~ 23 Å), we were surprised to observe such well resolved quantum dot states in our wider structure. In our experiment we have additional spatial resolution primarily due to the reduced area of the wire compared with the excitation spot. The wire excitons only radiate from a 70-Å-wide strip of the excitation spot, which effectively reduces the probe area by two orders of magnitude. We observe typically only ~ 2 peaks at any given position. This allows us to fully resolve all of the luminescent quantum dot states both spatially and spectrally.

We now discuss what we believe is evidence for an exciton relaxation bottleneck¹⁴. High-energy photo-excited excitons from the stem and arm QWs cascade into lower energy states in the quantum wire by emitting acoustic phonons in order to conserve energy. However, as the exciton localizes, the energy spacing between consecutive drops is predicted to grow rapidly. The

Figure 5 Magnification of the PL spectra from Fig. 4 at position 27.5 μm at several temperatures, showing the exciton relaxation bottleneck. As temperature is lowered, the PL peaks at $\sim 1,563.5$ and $\sim 1,562.0$ meV decrease in intensity, whereas the peak at $\sim 1,565.5$ meV increases in intensity. At the lowest temperature the exciton relaxation rate is so reduced that the excitons from the 1,565.5 meV peak radiate before relaxing into the ground state.

momentum of an emitted phonon is limited by the spatial extent of the localized exciton state. Because the energy of an acoustic phonon is determined by its momentum, an exciton relaxation bottleneck occurs when the exciton cannot emit a single phonon with enough energy to relax to the next level. Instead, the excitons must relax via higher ordered processes such as a multiphonon emission. The relaxation bottleneck has been controversial; it has to our knowledge never been directly observed, but it would be important as it could prevent the practical use of certain low-dimensional structures.

In Fig. 4 (20 K), there are several places along the wire that show unusually low-energy PL. In particular, Fig. 5 magnifies the set of states from position 27.5 μm of Fig. 4. Those states appear at precisely the same position to within our spatial resolution and are especially isolated from other states. The intensity of the PL peak at 1,565.5 meV increases substantially at lower temperatures, while the intensity of the lower energy peak at 1,563.6 meV actually decreases. The excitons apparently radiate from higher energy states before relaxing to the lowest available quantum dot state. This result could be interpreted in two ways. One possibility is that these are a set of isolated quantum dots, but so closely spaced that we cannot spatially separate them, in which case thermally activated excitons may hop between quantum dots. As the temperature is lowered, the excitons are unable to hop into the lowest states. However, given that the luminescent quantum dots are quite sparse along the wire and the island growth model implies that wide regions of the wire should be anticorrelated, such a series of closely spaced, isolated quantum dots is implausible. We prefer to interpret the series of states as excited states of a single quantum dot. The exciton relaxation bottleneck accounts for our observation that the lowest energy states are unoccupied. It is not obvious why the bottleneck vanishes at high temperatures, but multiphonon processes including emission and absorption are temperature-dependent. In this example the effect disappears above 15 K which corresponds to an acoustic phonon occupation up to 1.2 meV, comparable to the observed energy spacing between these states. This is a clear observation of a relaxation bottleneck in quantum dots. Although we have found other instances of this effect in our wire PL, Fig. 5 is the cleanest example that we have observed.

As previously mentioned, multiple quantum wires at T-intersections have been formed into an optically pumped laser⁴. Our quantum wire sample is nominally identical to one of the 22 matched quantum wires used in that laser. The laser in that report was operated at 2 K, at which according to our present measurements the individual quantum wires may have been dominated by quantum dot states at low exciton densities. It is an open question as to how or whether the quantum dot states reported here contribute to the lasing at high exciton densities. $\square$

Received 21 May; accepted 18 August 1997.

1. Loudon, R. One-dimensional hydrogen atom. *Am. J. Phys.* **27**, 649–655 (1959).
2. Chang, Y.-C., Chang, L. L. & Esaki, L. A new one-dimensional quantum well structure. *Appl. Phys. Lett.* **47**, 1324–1326 (1985).
3. Pfeiffer, L. N. *et al.* Formation of a high quality two-dimensional electron gas on cleaved GaAs. *Appl. Phys. Lett.* **56**, 1697–1699 (1990).
4. Wegscheider, W. *et al.* Lasing from excitons in quantum wires. *Phys. Rev. Lett.* **71**, 4071–4074 (1993).
5. Wegscheider, W., Pfeiffer, L. N., West, K. W. & Leibenguth, R. E. Current injection GaAs/AlGaAs quantum wire lasers fabricated by cleaved edge overgrowth. *Appl. Phys. Lett.* **65**, 2510–2512 (1994).
6. Goñi, A. R. *et al.* Observation of quantum wire formation at intersecting quantum wells. *Appl. Phys. Lett.* **61**, 1956–1958 (1992).
7. Someya, T., Akiyama, H. & Sakaki, H. Tightly confined one-dimensional states in T-shaped GaAs edge quantum wires with AlAs barriers. *Appl. Phys. Lett.* **66**, 3672–3673 (1995).
8. Hasen, J. *et al.* Large excitonic confinement in asymmetric quantum T wires. *Superlattices and Microstructures* (in the press).
9. Gislason, H., Langbein, W. & Hvam, J. M. Asymmetric GaAs/AlGaAs T wires with large confinement energies. *Appl. Phys. Lett.* **69**, 3248–3250 (1996).
10. Hess, H. F., Betzig, E., Harris, T. D., Pfeiffer, L. N. & West, K. W. Near-field spectroscopy of the quantum constituents of a luminescent system. *Science* **264**, 1740–1745 (1994).
11. Zrenner, A. *et al.* Quantum dots formed by interface fluctuations in AlAs/GaAs coupled quantum well structures. *Phys. Rev. Lett.* **72**, 3382–3385 (1994).
12. Brunner, K., Abstreiter, G., Böhm, G., Tränkle, G. & Weimann, G. Sharp-line photoluminescence of excitons localized at GaAs/AlGaAs quantum well inhomogeneities. *Appl. Phys. Lett.* **64**, 3320–3322 (1994).

13. Gammon, D., Snow, E. S., Shanabrook, B. V., Katzer, D. S. & Park, D. Fine structure splitting in the optical spectra of single GaAs quantum dots. *Phys. Rev. Lett.* **76**, 3005–3008 (1996).
14. Benisty, H., Sotomayor-Torres, C. M. & Weisbuch, C. Intrinsic mechanism for the poor luminescence properties of quantum-box systems. *Phys. Rev. B* **44**, 10945–10948 (1991).

Acknowledgements. We thank H. U. Baranger, T. D. Harris, P. B. Littlewood and R. Zimmermann for discussions.

Correspondence and requests for materials should be addressed to J.H. (e-mail: joel@lucent.com).

Large magnetoresistance in non-magnetic silver chalcogenides

R. Xu^{*†}, A. Husmann^{*}, T. F. Rosenbaum^{*}, M.-L. Saboungi[†], J. E. Enderby[†] & P. B. Littlewood[‡]

^{*} The James Franck Institute and Department of Physics, The University of Chicago, Chicago, Illinois 60637, USA

[†] Materials Science Division, Argonne National Laboratory, Argonne, Illinois 60439, USA

[‡] Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974, USA

Several materials have been identified over the past few years as promising candidates for the development of new generations of magnetoresistive devices. These range from artificially engineered magnetic multilayers¹ and granular alloys^{2,3}, in which the magnetic-field response of interfacial spins modulates electron transport to give rise to ‘giant’ magnetoresistance⁴, to the manganese perovskites^{5–7}, in which metal–insulator transitions driven by a magnetic field give rise to a ‘colossal’ magnetoresistive response (albeit at very high fields). Here we describe a hitherto unexplored class of magnetoresistive compounds, the silver chalcogenides. At high temperatures, the compounds Ag₂S, Ag₂Se and Ag₂Te are superionic conductors; below ~ 400 K, ion migration is effectively frozen and the compounds are non-magnetic semiconductors^{8,9} that exhibit no appreciable magnetoresistance¹⁰. We show that slightly altering the stoichiometry can lead to a marked increase in the magnetic response. At room temperature and in a magnetic field of ~ 55 kOe, Ag_{2+ δ} Se and Ag_{2+ δ} Te show resistance increases of up to 200%, which are comparable with the colossal-magnetoresistance materials. Moreover, the resistance of our most responsive samples exhibits an unusual linear dependence on magnetic field, indicating both a potentially useful response down to fields of practical importance and a peculiarly long length scale associated with the underlying mechanism.

Although substantial values of magnetoresistance are found for a variety of materials at low temperatures and in high magnetic fields, useful devices must operate at room temperature and in modest fields, $H < 1$ kOe. For non-magnetic metals, semimetals and semiconductors, the normalized magnetoresistance, $\Delta\rho(T, H)/\rho(T, 0)$, is usually negligible for such fields at $T \sim 300$ K (ref. 11). Narrow-gap semiconductors such as PbTe (ref. 12) and InSb (ref. 13) do indeed have a formidable low-field magnetoresistance at room temperature (0.01–1% for $H = 1$ kOe), but the characteristic resistivities are so large that the intrinsic device noise becomes the limiting factor. Balancing these constraints has led to permalloy (Ni_{0.81}Fe_{0.19}), with a saturation value $\Delta\rho/\rho \sim 2.5\%$ at $H \sim 5$ Oe, becoming the material of choice for standard applications¹⁴. Giant-magnetoresistance compounds seem to hold the greatest promise as replacement materials in the short term, with $\Delta\rho/\rho$ as high as 40% for $H \leq 1$ kOe being observed in Cu/Co multilayers at room temperature¹⁵. Both the high field scale and the resistivity ($\sim 10^4 \mu\Omega \text{ cm}$, comparable in magnitude with narrow-gap semiconductors) are present obstacles for potential colossal-magnetoresistance devices.

Silver selenide and silver telluride are narrow-gap self-doped degenerate n-type semiconductors. In their high-temperature (α) phase, electrical transport proceeds via a well-orchestrated coupling of microscopic lattice distortions to the ion migration dynamics. The structural transformation into the β -phase opens an energy gap as small as a few hundredths of an electronvolt. The electron density in the conduction band is determined largely by the stoichiometric index, δ . We cut specimens with typical dimensions 1–3 mm on a side and 1 mm thick from long boules with a designated purity of 99.999%. Slight gradients in stoichiometry along the length of the boule permitted the choice of samples with a range of electron densities. For both semiconductors, the nominal δ is approximately 0.01, with $\rho(300\text{ K}) \leq 10^3\ \mu\Omega\text{ cm}$. We performed four-probe resistivity and five-probe Hall coefficient measurements in fields up to 55 kOe using a conventional a.c. bridge technique in the ohmic and frequency-independent limits. The relative error in the magnetoresistivity is less than 0.01%; there is an absolute uncertainty, however, of 25% because of the finite extent of the silver epoxy contacts on millimetre-scale samples. We applied hydrostatic pressure with a BeCu piston–anvil self-locking pressure clamp, using silicone oil as the pressure medium and a chip of $(\text{V}_{0.99}\text{Ti}_{0.01})_2\text{O}_3$ as the manometer¹⁶.

We plot in Fig. 1 an overview of the temperature dependence of the resistivity of $\text{Ag}_{2+\delta}\text{Se}$ and $\text{Ag}_{2+\delta}\text{Te}$ for a series of applied magnetic fields. The field is applied perpendicular to the current direction; measurements of the longitudinal magnetoresistance also show a considerable effect, but at least a factor of 2 smaller than the transverse magnetoresistance. The temperature dependence of $\rho(H=0)$ is consistent with impurity band electrons on the delocalized side of the Mott metal–insulator transition¹⁷. Between $T = 4.2\text{ K}$ and $T = 60\text{ K}$ for $\text{Ag}_{2+\delta}\text{Se}$ (100 K for $\text{Ag}_{2+\delta}\text{Te}$), the resistivity rises because n is constant and the phonon scattering increases with T . At higher T , n becomes activated and ρ decreases with increasing T in accordance with the usual behaviour of small-gap semiconductors.

The large size and the magnetic-field dependence of the normalized magnetoresistance can be seen explicitly in Fig. 2. There is no evidence of saturation up to at least 55 kOe over the entire temperature range. $\Delta\rho/\rho$ increases linearly with increasing H at high T , with a tendency towards a slightly sublinear dependence

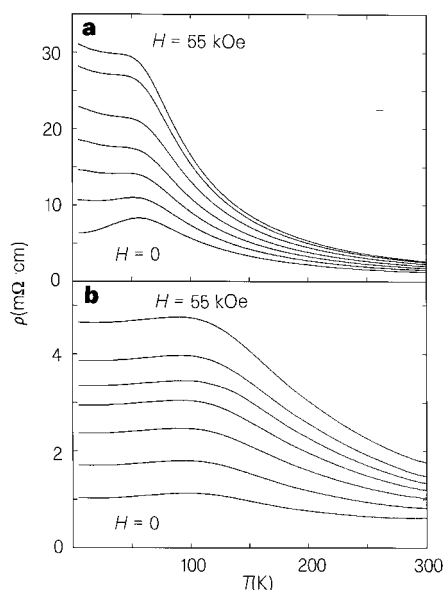


Figure 1 Resistivity ρ as a function of temperature T at a series of magnetic fields $H = 0, 9, 17, 26, 36, 48$ and 55 kOe for $\text{Ag}_{2+\delta}\text{Se}$ (**a**) and $H = 0, 10, 20, 29, 35, 43$ and 55 kOe for $\text{Ag}_{2+\delta}\text{Te}$ (**b**).

above tens of kOe at low T . The normal solution of the Boltzmann equation for magnetoresistance yields a dependence on H^2 , although scattering off the lattice and/or spin fluctuations can produce $\Delta\rho/\rho \propto H$. As can be seen in the inset to Fig. 2, the magnetoresistance falls below 1% by $H = 100\text{ Oe}$. The surprise in our data is that the pertinent field scale at which the magnetoresistance crosses over from a quadratic to a linear dependence on H is so small: $H_c \approx 10\text{ Oe}$, with the exact value being dependent on the range of fit. It is difficult to identify a physical length scale that corresponds to this magnetic field value. Yet the value of H_c seems to be related intimately to the large magnetoresistive response. Samples that exhibit a smaller magnetoresistance also display a larger crossover field, up to two orders of magnitude higher.

We compare in Fig. 3 the high-field (40 kOe), normalized magnetoresistance near room temperature for $\text{Ag}_{2+\delta}\text{Se}$, $\text{Ag}_{2+\delta}\text{Te}$, and a representative colossal-magnetoresistance compound, $\text{La}_{0.75}\text{Ca}_{0.25}\text{MnO}_3$ (ref. 18). The absolute value of the magnetoresistance is pronounced for all three, but of opposite sign—positive for the silver chalcogenides and negative for the manganite perovskites—reflecting completely different physical mechanisms. This difference is also exemplified in the contrasting temperature evolution of the magnetoresistance. $\Delta\rho/\rho$ of the colossal-magnetoresistance material changes rapidly in a narrow temperature range and must either be fine-tuned so that the metal–insulator transition temperature is near 300 K or be modulated via interfacial tunnelling in fine-grained polycrystals¹⁹. In contrast, the silver chalcogenide data evolve smoothly with temperature and exhibit no sharp features.

The question remains of how to optimize the normalized magnetoresistance for $\text{Ag}_{2+\delta}\text{Se}$ and $\text{Ag}_{2+\delta}\text{Te}$. As demonstrated in Fig. 4, it seems that the carrier density is a key variable. Analysis of the linear Hall response at $H = 5\text{ kOe}$ in the single carrier ‘exhaustion’ regime (temperatures below the maximum of $\rho(T)$ in Fig. 1, where the Hall coefficient, R_H , becomes temperature-independent), shows an increasing magnetoresistance with decreasing R_H (increasing n). An increase in the intrinsic carrier density from 2.5×10^{17} to $9 \times 10^{17}\text{ cm}^{-3}$ results in an increase in $\Delta\rho(T = 4.5\text{ K}, H = 55\text{ kOe})/\rho(T = 4.5\text{ K}, H = 0)$ from 200% to 350%. Moreover, identically doped samples of $\text{Ag}_{2+\delta}\text{Se}$ and $\text{Ag}_{2+\delta}\text{Te}$ have identically large magnetoresistances.

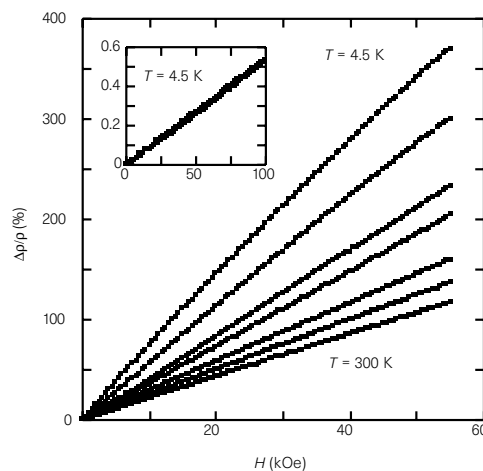


Figure 2 The normalized magnetoresistance $\Delta\rho(T, H)/\rho(T, 0)$ for $\text{Ag}_{2+\delta}\text{Se}$ as a function of magnetic field H at a series of temperatures $T = 4.5, 30, 60, 90, 180, 270$ and 300 K . Inset, the linear field dependence continues down to H values of a few Oe, implying a very large characteristic length scale.

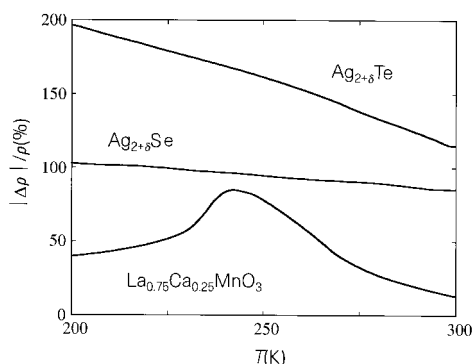


Figure 3 Comparison of the normalized magnetoresistance of the silver chalcogenides and a representative colossal-magnetoresistance material at $H = 40$ kOe.

The classical, orbital magnetoresistance is controlled by the ratio of the cyclotron frequency to the scattering rate, $\omega_c \tau = eH\tau/m^*c$. Doped narrow-gap semiconductors have small effective masses, $m^* \approx 0.02m_e$ ($0.08m_e$) for Ag_2Te (Ag_2Se) (ref. 20), which gives a large magnetoresistance. Furthermore, good metallic conduction can be obtained even at low carrier densities because the combination of a small m^* and a large dielectric constant pushes the metal–insulator transition boundary to low densities. One expects a positive magnetoresistance, quadratic at low fields, and saturating for $\omega_c \tau > 1$. A non-saturating, but still quadratic, magnetoresistance can occur with perfect compensation along with a vanishing Hall effect, as would be expected in the intrinsic regime at high T .

Despite the advantageous standing of narrow-gap semiconductors, the classical picture seems incapable of explaining the silver chalcogenide data, even allowing for a crossover from degenerate to non-degenerate transport seen in $\rho(T > 100 \text{ K})$. The magnetoresistance is linear down to very low fields; it shows little sign of saturation at high fields, even though $\omega_c \tau > 5$ at the highest fields for the $\text{Ag}_{2+\delta}\text{Te}$ sample in Fig. 1; and it is only weakly temperature-dependent, with no evidence for a crossover associated with degenerate to non-degenerate carriers. Inhomogeneous density fluctuations, on a scale larger than the mean free path but much smaller than the sample size, can produce a magnetoresistance linear in H with $\Delta\rho/\rho \approx \omega_c \tau$ (ref. 21). The electronic and point defects that occur in the silver chalcogenides for $\delta \geq 10^{-4}$ (refs 22, 23) might provide a source of such density fluctuations, leading in principle to a high-field transverse magnetoresistance of the observed size. However, this mechanism cannot explain either the anomalously low H_c (at $\omega_c \tau \ll 1$) or the correlation between R_H and $\Delta\rho/\rho$ demonstrated in Fig. 4.

We have studied samples doped as far away from 2:1 stoichiometry as $\text{Ag}_{0.70}\text{Se}_{0.30}$, where $R_H(T \rightarrow 0)$ approaches $0.4 \text{ cm}^3 \text{ C}^{-1}$ ($n > 10^{19} \text{ cm}^{-3}$) and $\rho(T = 300 \text{ K}) = 350 \mu\Omega \text{ cm}$ (Fig. 4). In this case, $\Delta\rho/\rho$ falls to 110%, implying an optimal value for δ between 0.01 and 0.33. A significant advantage of the silver chalcogenides is that the Ag stoichiometry could be controlled precisely in future experiments through electrochemical techniques because silver β -alumina is readily available. The results of applying hydrostatic pressure are also encouraging. A pressure of only 7 kbar halved the room-temperature resistivity of the $\text{Ag}_{2+\delta}\text{Se}$ sample of Fig. 2 without reducing the magnetoresistance. In contrast, the magnetoresistance in colossal-magnetoresistance compounds is severely suppressed by pressure²⁴. Hence, the strategy of introducing internal pressure through chemical means or, alternatively, interfacial stress in a

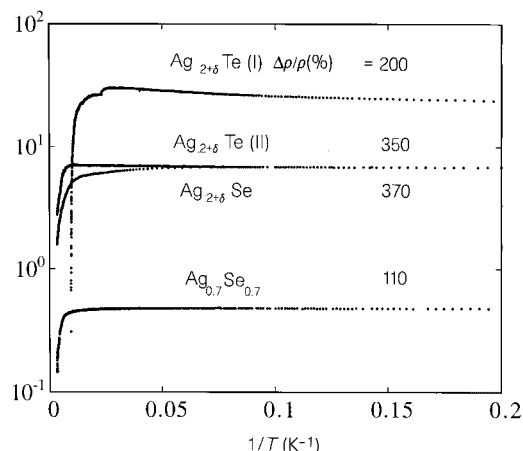


Figure 4 Hall coefficient R_H (n-type) plotted against T^{-1} for a set of silver chalcogenide samples labelled by their magnetoresistive response (at $T = 4.5 \text{ K}$ and $H = 55 \text{ kOe}$). Carrier density in the intrinsic, low-temperature limit seems to be a key causal parameter.

silver chalcogenide thin film, may help satisfy the device goal of a small ρ without sacrificing the large magnetoresistance. □

Methods

High-purity Ag_2Se and Ag_2Te (99.999% pure) were used as received from Johnson Matthey Corporation and United Mineral Corporation. The stoichiometric materials were ground and loaded into four outgassed fused silica ampoules inside a helium glovebox, and appropriate amounts of silver or selenium were added to reach the desired compositions. The selenium was treated to remove the oxide coatings by filtering through a fine-size frit near the melting point. The samples were sealed under vacuum, heated in a rocking furnace above their respective melting points for 24 h and left to cool in a horizontal position. Powder X-ray diffraction gave the expected orthorhombic structure at room temperature with a grain size of at least 500 nm. No additional phases were detected in any sample. SQUID magnetometry on $\text{Ag}_{2+\delta}\text{Se}$ showed no evidence for paramagnetic impurities with an essentially temperature-independent diamagnetic response, $\chi(T = 4.5 \text{ K}) = -1.7 \times 10^{-7} \text{ cm}^{-3}$.

Received 9 May; accepted 27 August 1997.

- Baibich, M. *et al.* Giant magnetoresistance of (001)/Fe/(001)Cr magnetic superlattices. *Phys. Rev. Lett.* **61**, 2472–2475 (1988).
- Berkowitz, A. E. *et al.* Giant magnetoresistance in heterogeneous Cu–Co alloys. *Phys. Rev. Lett.* **68**, 3745–3748 (1992).
- Xiao, J. Q., Jiang, J. S. & Chien, C. L. Giant magnetoresistance in nonmultilayer magnetic systems. *Phys. Rev. Lett.* **68**, 3749–3752 (1992).
- Camley, R. E. & Stamps, R. L. Magnetic multilayers: spin configurations, excitations and giant magnetoresistance. *J. Phys. Cond. Mater.* **5**, 3727–3786 (1993).
- Searle, C. W. & Wang, S. T. Studies of the ionic ferromagnet (LaPb)MnO₃. III. Ferromagnetic resonance studies. *Can. J. Phys.* **47**, 2703–2708 (1969).
- Jin, S. *et al.* Thousandfold change in resistivity in magnetoresistive La–Ca–Mn–O films. *Science* **264**, 413–415 (1994).
- Tokura, Y. *et al.* Origins of colossal magnetoresistance in perovskite-type manganese oxides. *J. Appl. Phys.* **79**, 5288–5291 (1996).
- Dalven, R. & Gill, R. Energy gap in β - Ag_2Se . *Phys. Rev.* **159**, 645–649 (1967).
- Junod, P., Heidiger, H., Kilchör, B. & Wulschlegel, J. Metal–non-metal transition in silver chalcogenides. *Phil. Mag.* **36**, 941–958 (1977).
- Junod, P. Relations entre la structure cristalline et les propriétés électroniques des combinaisons Ag_2S , Ag_2Se , Cu_2Se . *Helv. Phys. Acta* **32**, 567–600 (1959).
- Jan, J.-P. in *Solid State Physics*, Vol. 5 (eds Seitz, F. & Turnbull, D.) 1–96 (Academic, New York, 1957).
- Allgaier, R. S., Restorff, J. B. & Houston, B. Weak- and strong-field magnetoresistance in (111)-oriented n-type PbTe epitaxial films between 1.8 and 300 K. *J. Appl. Phys.* **53**, 3110–3116 (1982).
- Schönwald, H. Die Beweglichkeit der langsamen und schnellen Löcher in Indiumantimonid. *Z. Naturf.* **19a**, 1276–1296 (1964).
- Brug, J. A., Anthony, T. C. & Nickel, J. H. Magnetic recording head materials. *Mater. Res. Soc. Bull.* **21**, 23–27 (1996).
- Parkin, S. S. P., Li, Z. G. & Smith, D. J. Giant magnetoresistance in antiferromagnetic Co/Cu multilayers. *Appl. Phys. Lett.* **58**, 2710–2712 (1991).
- Rosenbaum, T. F., Carter, S. A. & Honig, J. M. High sensitivity sensor for moderate pressures. *Rev. Sci. Instrum.* **67**, 617–618 (1996).
- Mott, N. F. *The Metal–Insulator Transition* (Taylor and Francis, London, 1974).
- Schiffer, P. *et al.* Low temperature magnetoresistance and the magnetic phase diagram of $\text{La}_{1-x}\text{Ca}_x\text{MnO}_3$. *Phys. Rev. Lett.* **75**, 3336–3339 (1995).

19. Hwang, H. *et al.* Spin-polarized intergrain tunneling in $\text{La}_{2/3}\text{Sr}_{1/3}\text{MnO}_3$. *Phys. Rev. Lett.* **77**, 2041–2044 (1996).
20. Aliev, S. A. & Aliev, F. F. Band parameters and energy structure of $\beta\text{-Ag}_2\text{Se}$. *Izv. Akad. Nauk SSSR, Neorgan. Mater.* **21**, 1869–1872 (1985).
21. Herring, C. Effect of random inhomogeneities on electrical and galvanomagnetic measurements. *J. Appl. Phys.* **31**, 1939–1953 (1960).
22. v. Oehsen, U. & Schmalzried, H. Thermodynamic investigations of Ag_2Se . *Ber. Bunsenges. Phys. Chem.* **85**, 7–14 (1981).
23. Valverde, N. Coulometrische Titrationen zur Bestimmung des Homogenitätsbereiches von festem Silbersulfid, Silberselenid und Silbertellurid. *Z. Phys. Chem. NF* **70**, 113–127 (1970).
24. Khazeni, K. *et al.* Effect of pressure on the magnetoresistance of single crystal $\text{Nd}_{0.5}\text{Sr}_{0.36}\text{Pb}_{0.14}\text{MnO}_{3-\delta}$. *Phys. Rev. Lett.* **76**, 295–298 (1996).

Acknowledgements. We thank D. L. Price and B. J. Wuensch for discussions. The work at the University of Chicago was supported primarily by the MRSEC Program of the National Science Foundation. The work at Argonne National Laboratory was supported by DOE Basic Energy Sciences.

Correspondence and requests for materials should be addressed to T.F.R. (e-mail: t-rosenbaum@uchicago.edu).

NMR evidence for excess non-bridging oxygen in an aluminosilicate glass

Jonathan F. Stebbins* & Zhi Xu†

* Department of Geological and Environmental Sciences, and

† Department of Materials Science and Engineering, Stanford University, Stanford, California 94305-2115, USA

The most common of man-made glasses have aluminosilicate compositions, and such glasses also form from rapidly cooling magmas¹. Oxygen is the most abundant element in these materials, where it occupies either ‘bridging’ (BO) or ‘non-bridging’ (NBO) sites. BOs link two AlO_4 or SiO_4 tetrahedra, thereby providing strong, long-lived bonds between the smallest structural units of the aluminosilicate network. NBOs provide a relatively weak connection between one tetrahedral cation (Al or Si) and one or more network modifier cations—such as Ca^{2+} or Na^+ —that are not an integral part of the tetrahedral network. The relative abundance of these weakly bonded NBOs is critical in determining the thermodynamic and dynamical properties of aluminosilicate glasses and melts^{1–3}. For glasses of ‘tectosilicate’ composition, where the charge of the modifier cation equals the number of aluminium atoms (as in $\text{NaAlSi}_3\text{O}_8$ or $\text{CaAl}_2\text{Si}_2\text{O}_8$), the conventional view of glass structure is that only BOs are present^{1,4}. Here we present experimental observations that contradict this view. Our NMR measurements of $\text{CaAl}_2\text{Si}_2\text{O}_8$, which determine directly the relative abundances of BO and NBO, indicate that a considerable amount of NBO can be present in a tectosilicate glass. These excess NBOs will increase the entropy and heat capacity of the corresponding liquid and decrease its viscosity, as well as modifying flow and diffusion mechanisms^{2,3}. As the most common rhyolitic magmas and the molten precursors of glass ceramics have near-tectosilicate compositions^{1,4}, our results require a reassessment of the high-temperature liquid properties that control many processes in the Earth and in industry.

The viscosity of a binary silicate liquid (such as CaSiO_3 or Na_2SiO_3) increases progressively as the network modifier oxide (CaO or Na_2O) is replaced by Al_2O_3 . At a point near the 1:1 ratio of CaO and Al_2O_3 (or Na_2O and Al_2O_3), the viscosity reaches a maximum⁵. According to the conventional model of aluminosilicate structure^{1,4}, this viscosity increase arises from a gradual conversion of NBO to BO. This follows from requiring that all of the Al remain in AlO_4 tetrahedra, each oxygen atom of which is bonded to no more than one other tetrahedral cation. Because this replacement results in the substitution of relatively weak bonds between modifier cations and NBO by stronger, long-lived bonds between BO and tetrahedral cations, the viscosity increases. Furthermore, with increasing Al_2O_3 content, fewer modifier cations are available to

attack and weaken the strong BO bonds, thereby further contributing to the increase in viscosity. This general picture provides a good qualitative explanation for the behaviour of complex natural systems, such as the enormous difference in viscosity—and thus eruptive style—between basaltic and rhyolitic magmas. Basalts have relatively low silica contents, on average about 1 NBO per tetrahedral cation (T), and usually erupt as quiescent flows, whereas rhyolites have much higher silica contents, NBO/T close to 0, and often erupt as viscous domes or in catastrophic explosions.

At a 1:1 ratio of the charge of the modifier cation to the number of aluminium atoms, the liquid has the stoichiometry of tectosilicates such as feldspars and quartz. Such liquids and glasses are often described as having a ‘fully polymerized’ structure—a somewhat randomized version of the corresponding crystal structure, with no NBO. The microscopic confirmation of this assumption is limited, in spite of the fact that it is the basis for a number of models describing the viscosity⁶ and thermodynamics⁷ of aluminosilicates. X-ray scattering data on feldspar-composition glasses are consistent with disordered framework structures⁸ as expected for the ‘fully polymerized’ liquids and glasses, but minor amounts of other non-framework species, such as NBO or AlO_5 , generally cannot be detected. In Raman^{1,9} and ^{29}Si MAS NMR spectra^{10,11}, peaks for tetrahedral units with NBO overlap those containing only BO, making it also difficult to detect variations from the conventional approximation.

Precise measurements of the viscosity of Na- and Ca-aluminosilicate liquids have demonstrated that the viscosity maxima actually occur at a higher Al_2O_3 content than the ideal 1:1 ratio, suggesting that there is a considerable concentration of NBO even at the tectosilicate composition^{12,13}. Displaced viscosity maxima in Ca-aluminosilicates were also suggested by less accurate, earlier data⁶. Oxygen-17 NMR is an ideal method by which to investigate directly whether NBO is indeed present at the 1:1 ratio^{14,15}. Figure 1 shows ^{17}O magic-angle-spinning (MAS) NMR spectra of two glass samples

Table 1 Composition (in mol%) of glass samples similar to $\text{CaAl}_2\text{Si}_2\text{O}_8$

	Glass A	Glass B
SiO_2	49.6 ± 0.1	49.4
Al_2O_3	25.0 ± 0.2	25.2
CaO	25.5 ± 0.1	25.3
NBO predicted	0.5 ± 0.2	0.1
NBO observed	5 ± 1	4

Glass A was made by melting dried CaO , Al_2O_3 and 45% ^{17}O SiO_2 at $1,610^\circ\text{C}$ in a sealed platinum tube, and was analysed by electron microprobe. Uncertainties in composition are 2σ estimates from 4 to 8 repeated analyses. Glass B was made from glass A by adding 0.50 wt% Al_2O_3 and remelting. Both glasses contain about 0.07 mol% CoO , added to reduce spin-lattice relaxation time. NBO content (% of total oxygen) is predicted from the excess of $(\text{CaO} + \text{CoO})$ over Al_2O_3 . The observed NBO content was derived from the 1D MAS NMR spectra assuming a symmetric peak shape, as is commonly observed in binary silicate glass compositions^{15,16}.

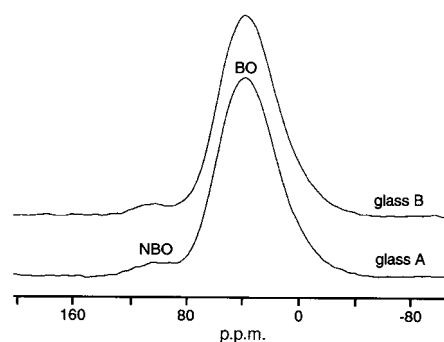


Figure 1 Oxygen-17 MAS NMR spectra of glasses A and B. The spectra were referenced to H_2O and collected at 54.2 MHz with a spinning rate of 12 kHz. Single $0.3\text{-}\mu\text{s}$ pulses (with roughly 10° radiofrequency tip angle for solid) were used, with a delay of 1 s. No differential relaxation was observed.

with compositions close to that of $\text{CaAl}_2\text{Si}_2\text{O}_8$ (Table 1). The small peak near 100 p.p.m. is unambiguously attributable to NBO sites: it is close to the position of the NBO peak in CaO-SiO_2 binary glasses¹⁶⁻¹⁸ and is distinct from all known BO positions. (It is *not* in the correct position to be a spinning sideband from either the central or the satellite transitions, including the centre sideband for the latter.) A control experiment confirmed that the NBO peak is not due to sample heterogeneity or to excess CaO : addition of 0.5 wt % Al_2O_3 to glass A and remelting yielded glass B (Table 1) which also displays a peak at the NBO position (Fig. 1). The high NBO content cannot be due to dissolved water, as the glass composition is known to be non-hygroscopic. Furthermore, the glass was melted at a temperature above 1,600 °C, where water solubility is very low. Both the NBO and BO peaks have featureless shapes typical of alkaline earth silicates^{16,18}. The observed concentration of NBO (4 to 5%, Table 1) is about an order of magnitude higher than predicted by the standard model, even when allowing for stoichiometric variations due to experimental error in composition (<0.7%).

Figure 2 shows the ^{17}O triple quantum MAS (3QMAS) NMR spectrum for glass B. This new technique reveals additional structural detail by eliminating second-order quadrupolar broadening in

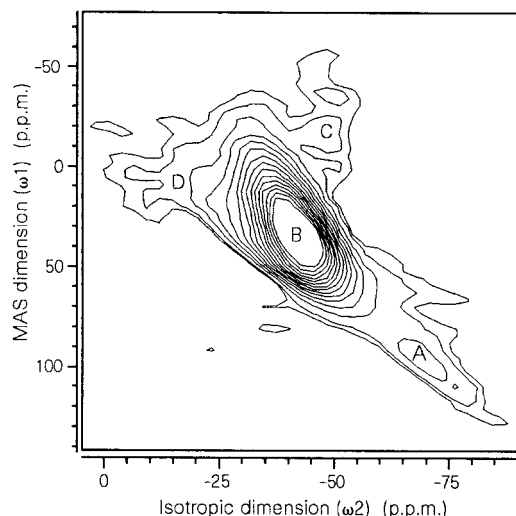


Figure 2 Oxygen-17 3QMAS NMR spectrum of glass B. The spectrum was collected with a two-pulse sequence (180° - 360°)²⁸ and a delay of three times the measured relaxation time of about 5 s. We collected 480 transients per time increment. Other parameters were as in Fig. 1. Data were processed as previously described²⁰. Contours are drawn at 3, 5, 10, 17, ..., 70% of highest peak intensity. Spectra for glass A are nearly identical.

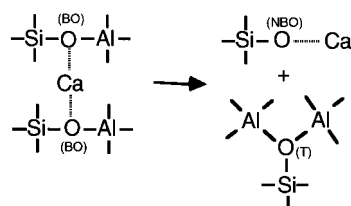


Figure 3 Cartoon of transformation of bridging oxygen (BO) to non-bridging oxygen (NBO). Incompletely drawn bonds are all to O. If, as shown, all Al remains as AlO_4 , the new structural unit formed is considered to be a 'tricluster'; alternatively, if AlO_5 or AlO_6 sites are formed, the 'tricluster' oxygen (T) may itself behave as an NBO. An OAl_3 tricluster could also form, if one of the initial BO is Al-O-Al instead of Si-O-Al .

the second or 'isotropic' dimension^{19,20}. The structurally most relevant parameters of isotropic chemical shift (δ_{CS} , in p.p.m.) and quadrupolar coupling constant (C_q , in MHz) can be calculated from peak positions in the two dimensions, assuming that the quadrupolar asymmetry parameter (η_q) is within the typical range of 0 to 0.5 (ref. 20). In the spectrum, the NBO peak (A) is clearly defined (113 ± 2 p.p.m., 2.9 ± 0.2 MHz). An NBO peak was not observed in a recent ^{17}O 3QMAS spectrum of $\text{NaAlSi}_3\text{O}_8$ glass²¹, perhaps because the concentration was lower, or, more probably, because of the overlap of NBO and BO peaks for Na silicates^{15,22}.

Besides the NBO peak and the main Si-O-Al peak (B) (61 p.p.m., 3.5 MHz; the large isotropic breadth indicates considerable disorder) two other features are visible. Peak C (50 p.p.m., 5.5 MHz) is in a position similar to Si-O-Si peaks in crystalline aluminosilicates^{23,24} as well as in $\text{NaAlSi}_3\text{O}_8$ glass²¹, with some shift due to coordination by Ca instead of Na. Its presence may be the result of considerable Si/Al disorder, with some Si-O-Si and Al-O-Al linkages instead of the strict Si-O-Al alternation found in the crystal. The Al-O-Al links, in turn, may be promoted by the relatively high held strength of Ca^{2+} , which favours negative charge concentration^{3,10,25}. For this reason, formation of NBOs is more likely in Ca-containing glasses, and may, at least partly, explain the lower viscosity of liquid $\text{CaAl}_2\text{Si}_2\text{O}_8$ compared with liquid NaAlSiO_4 (ref. 2). The Al-O-Al peak that should accompany Al/Si disorder in the $\text{CaAl}_2\text{Si}_2\text{O}_8$ glass may be hidden under a shoulder of the Si-O-Al peak, or could be absent if some other Al-rich species is formed, as suggested below and in Figure 3.

Feature D (20 p.p.m., 2.3 MHz) may be the result of the new structural unit that must be created when NBO is present near the tectosilicate composition. Aluminium-27 3QMAS NMR data for $\text{CaAl}_2\text{Si}_2\text{O}_8$ (ref. 20) indicate that the concentrations of AlO_5 or AlO_6 units are below 1 or 2%, and likely candidates for the new species are therefore OAl_3 or OAl_2Si triclusters, in which Al remains in four-coordination (Fig. 3). Toplis *et al.*^{12,13} suggested that a reaction to form NBO and triclusters could explain the displaced viscosity maxima in aluminosilicate liquids; we conclude here that this general idea is indeed likely to be correct.

Our findings strongly suggest that simple pictures of aluminosilicate melt speciation, which include only the structural units required by stoichiometry, need to be fundamentally revised. Despite the small total concentrations involved, the presence of NBO in $\text{CaAl}_2\text{Si}_2\text{O}_8$ glass nevertheless has important implications for models of the transport properties and thermodynamic activities of components containing NBO (such as chain and sheet silicates). NBO sites can be considered as weak links in the silicate network, and their abundance is thus probably the most important single factor in controlling viscosity in aluminosilicate liquids. In tectosilicate compositions, the hypothesized absence of NBO has required the postulation of mechanisms for viscous flow that are different from those in liquids richer in network modifiers. For example, the formation of transient SiO_5 (and/or AlO_5) groups (which remain at very low concentrations) from NBO and SiO_4 (or AlO_4) groups has been suggested as a key mechanism in viscous flow in non-framework compositions^{3,26}—a model that cannot be applied to geologically and technologically important tectosilicate liquids if NBO is assumed to be absent. The close relationship of this mechanism to that suggested for the pressure-induced formation of SiO_5 groups^{3,27} suggests that minor amounts of NBO could also play an important role in the densification of tectosilicate liquids at high pressure. □

Received 20 May; accepted 8 September 1997.

1. Mysen, B. O. *Structure and Properties of Silicate Melts* (Elsevier, Amsterdam, 1988).
2. Richet, P. & Neuville, D. R. in *Thermodynamic Data* (ed. Saxena, S. K.) (Springer, New York, 1992).
3. Stebbins, J. F. in *Structure, Dynamics, and Properties of Silicate Melts* (eds Stebbins, J. F., McMillan, P. F. & Dingwell, D. B.) (Mineralogical Society of America, Washington DC, 1995).
4. Navrotsky, A. *Physics and Chemistry of Earth Materials* (Cambridge Univ. Press, 1994).
5. Riebling, E. F. Structure of sodium aluminosilicate melts containing at least 50 mol% SiO_2 at 1500 °C. *J. Chem. Phys.* **44**, 2857-2865 (1966).

6. Bottinga, Y. & Weill, D. F. The viscosity of magmatic silicate liquids: a model for calculation. *Am. J. Sci.* **272**, 438–475 (1972).
7. Weill, D. F., Hon, R. & Navrotsky, A. in *Physics of Magmatic Processes* (ed. Hargraves, R. B.) (Princeton University Press, Princeton, NJ, 1980).
8. Taylor, M. & Brown, G. E. Jr Structure of mineral glasses—I. The feldspar glasses NaAlSi₃O₈, KAlSi₃O₈, CaAl₂Si₂O₈. *Geochim. Cosmochim. Acta* **43**, 61–75 (1979).
9. Mysen, B. O., Virgo, D. & Seifert, F. A. The structure of silicate melts: implications for chemical and physical properties of natural magma. *Rev. Geophys. Space Phys.* **20**, 353–383 (1982).
10. Murdoch, J. B., Stebbins, J. F. & Carmichael, I. S. E. High-resolution ²⁹Si NMR study of silicate and aluminosilicate glasses: the effect of network-modifying cations. *Am. Mineral.* **70**, 332–343 (1985).
11. Merzbacher, C. I., Sherriff, B. L., Hartman, J. S. & White, W. B. A high resolution ²⁹Si and ²⁷Al NMR study of alkaline earth aluminosilicate glasses. *J. Non-Cryst. Solids* **124**, 194–206 (1990).
12. Toplis, M. J. & Dingwell, D. B. Viscosity maxima of melts close to the 'charge balanced' join in the systems (Na₂O, CaO, MgO)–Al₂O₃–SiO₂: implications for the structural role of aluminium. *Eos Trans. Am. Geophys. Un.* **77**, F848 (1996).
13. Toplis, M. J., Dingwell, D. B. & Lenci, T. Peraluminous viscosity maxima in Na₂O–Al₂O₃–SiO₂ liquids: the role of triclusters in tectosilicate melts. *Geochim. Cosmochim. Acta* **61**, 2605–2612 (1997).
14. Farnan, I. *et al.* Quantification of the disorder in network modified silicate glasses. *Nature* **358**, 31–35 (1992).
15. Florian, P., Vermillion, K. E., Grandinetti, P. J., Farnan, I. & Stebbins, J. F. Cation distribution in mixed alkali disilicate glasses. *J. Am. Chem. Soc.* **118**, 3493–3497 (1996).
16. Kirkpatrick, R. J. in *Spectroscopic Methods in Mineralogy and Geology* (ed. Hawthorne, F. C.) (Mineralogical Society of America, Washington DC, 1988).
17. Oglesby, J. V., Xu, Z. & Stebbins, J. F. Non-bridging oxygen order and disorder in mixed alkaline earth silicate glasses. *Eos Trans. Am. Geophys. Un.* **77**, F834 (1996).
18. Stebbins, J. F., Oglesby, J. V. & Xu, Z. Disorder among network modifier cations in silicate glasses: new constraints from triple-quantum oxygen-17 NMR. *Am. Mineral.* (in the press).
19. Frysman, L. & Harwood, J. S. Isotropic spectra of half-integer quadrupolar spins from bidimensional magic-angle spinning NMR. *J. Am. Chem. Soc.* **117**, 5367–5368 (1995).
20. Baltisberger, J. H., Xu, Z., Stebbins, J. F., Wang, S. & Pines, A. Triple-quantum two-dimensional ²⁷Al magic-angle spinning nuclear magnetic resonance spectroscopic study of aluminosilicate and aluminate crystals and glasses. *J. Am. Chem. Soc.* **118**, 7209–7214 (1996).
21. Dirken, P. J., Kohn, S. C., Smith, M. E. & van Eck, E. R. H. Complete resolution of Si–O–Si and Si–O–Al fragments in an aluminosilicate glass by ¹⁷O multiple quantum magic angle spinning NMR spectroscopy. *Chem. Phys. Lett.* **266**, 568–574 (1997).
22. Xue, X., Stebbins, J. F. & Kanzaki, M. Correlations between O-17 NMR parameters and local structure around oxygen in high-pressure silicates and the structure of silicate melts at high pressure. *Am. Mineral.* **79**, 31–42 (1994).
23. Xu, Z. & Stebbins, J. F. NMR study of oxygen isotope exchange in stilbite. *Eos Trans. Am. Geophys. Un.* **77**, F827 (1996).
24. Xu, Z. & Stebbins, J. F. Oxygen sites in the zeolite, stilbite: a comparison of static, MAS, VAS, DAS, and triple quantum MAS NMR techniques. *Solid State NMR* (in the press).
25. McMillan, P. A. Raman study of glasses in the system CaO–MgO–SiO₂. *Am. Mineral.* **69**, 645–659 (1984).
26. Farnan, I. & Stebbins, J. F. A high temperature ²⁹Si NMR investigation of solid and molten silicates. *J. Am. Chem. Soc.* **112**, 32–39 (1990).
27. Xue, X., Stebbins, J. F., Kanzaki, M., McMillan, P. F. & Poe, B. Pressure-induced silicon coordination: NMR, Raman, and infrared spectroscopy. *Am. Mineral.* **76**, 8–26 (1991).
28. Massiot, D. *et al.* Two-dimensional magic-angle spinning isotropic reconstruction sequences for quadrupolar nuclei. *Solid State NMR* **6**, 73–84 (1996).

Acknowledgements. We thank R. Jones (Stanford Centre for Materials Research) for the microprobe analyses. This work was funded by the US National Science Foundation.

Correspondence and requests for materials should be addressed to J.F.S. (e-mail: stebbins@pangea.stanford.edu).

Observation of stratospheric ozone depletion in rocket exhaust plumes

M. N. Ross*, J. R. Benbrook†, W. R. Sheldon‡, P. F. Zittel‡ & D. L. McKenzie‡

* Environmental Systems Directorate, The Aerospace Corporation, PO Box 92957, Los Angeles, California 90009, USA

† Atmospheric and Space Physics Group, University of Houston, Houston, Texas 77204, USA

‡ Space and Environment Technology Center, The Aerospace Corporation, PO Box 92957, Los Angeles, California 90009, USA

Although modelling studies have predicted^{1–7} that particulate and reactive gas-phase species in the exhaust plume of large rockets might cause significant local ozone depletion, the actual response of the stratosphere after rocket launches has never been directly determined. Here we report comprehensive measurements that follow the evolution of stratospheric ozone in the wake of two Titan IV rockets launched on 12 May and 20 December 1996. In both cases, ozone concentrations dropped to near-zero values in

the plume wake, across regions four to eight kilometres wide, within 30 minutes after launch; intense ozone loss persisted for 30 minutes after which time concentrations recovered to ambient levels. Our data indicate that the number of ozone molecules lost in the plume regions significantly exceed the number of chlorine molecules deposited by the two rockets. This suggests that a catalytic cycle based on Cl₂O₂, other than Cl₂, and unique to solid rocket motor (SRM) plumes might be responsible for our observations. However, the limited spatial and temporal extent of the observed ozone losses implies that neither the catalytic Cl₂O₂ cycle nor other reactions involving exhaust compounds from large solid-fuelled rockets have a globally significant impact on stratospheric chemistry.

Although rocket motor emissions represent a small fraction of the total anthropogenic impact on stratospheric chemistry, prudence requires a careful evaluation of this impact, particularly on stratospheric ozone^{8,9}. Increasingly sophisticated modelling efforts have converged on the view that the present fleet of solid-fuelled rockets contributes only negligibly to ozone depletion on a global scale¹⁰ but there is less consensus on the magnitude of local effects. Predictions of the extent of local transient ozone loss in the expanding exhaust plume of a Space Shuttle or Titan IV rocket vary from a few tens of per cent⁴ to 100% (ref. 7) over horizontal distances of several kilometres.

The first direct observations of the effects of rocket launches have been reported by Pergament *et al.*¹¹, who measured a reduction in ozone levels of 40% during a single plume pass at an altitude of 18 km, 13 min after a 1975 Titan III launch. The observed loss suggests the presence of an ozone-destroying exhaust component, but the measurement, of uncertain reliability, was never repeated.

Table 1 Details of the WB-57F plume passes through the Titan IV exhaust plumes on 12 May and 20 December 1996

Pass number	Time after launch (min)	Pass length (km)	Ozone loss (%)
12 May 1996			
1	16	1.8	34
2	21	1.8	54
3	27	1.8	67
4	29	3.5	66
5	34	2.6	98
6	38	2.0	0
7	43	4	53
8	49	6.6	78
9	55	8.4	73
10	63	8.2	34
11	71	12.2	0
12	82	10.0	0
13	91	11.0	12
14	104	8.4	0
15	112	3.9	19
16	114	5.8	0
17	121	4.4	53
18	130	8.6	0
20 December 1996			
1	20	2.0	0
2	22	2.4	55
3	26	4.0	72
4	30	4.2	88
5	34	4.0	81
6	45	2.8	71
7	46	2.0	60
8	51	4.8	86
9	56	3.2	74
10	60	6.8	37
11	65	5.4	37
12	81	5.6	33
13	87	2.0	42

Time after launch represents the beginning of the pass, to the nearest minute. Ozone loss is based on the mean value of ozone concentration in the plume calculated from all operating channels of the ultraviolet photometers. For the 12 May data, a report of 0% ozone loss indicates either that ozone concentration in the plume was not different from ambient, or that the aircraft did not actually intercept the plume, as reported by the WB-57F pilot. This ambiguity is not a factor in the 20 December data.

Analysis of ozone data from the Total Ozone Mapping Spectrometer (TOMS) carried on the Nimbus 7 spacecraft, in contrast, failed to provide evidence for rapid local decreases in ozone after several Space Shuttle launches¹². However, no firm conclusions can be drawn from the TOMS data either with respect to the Space Shuttle or rocket exhaust effects in general, because the complex aerosol and gas environment of rocket exhaust plumes might give rise to anomalous scattering and absorption which, in turn, would compromise the TOMS data¹³.

As part of a United States Air Force programme to measure the local response of stratospheric ozone to rocket emissions, a NASA WB-57F aircraft carried two nearly identical ultraviolet absorption photometers into the wakes of two Titan IV rockets on 12 May and 20 December 1996. The two instruments, carried in the WB-57F nose payload bay, had previously provided reliable and accurate ozone measurements in the stratosphere on various rocket and balloon platforms^{14,15}. The Titan IVs—propelled by solid-fuelled rocket motors (SRMs)—were launched from Vandenberg Air Force Base (120° 37' W, 34° 48' N) at 13:33 Pacific Standard Time (PST) and 10:04 PST, respectively. In each case, the WB-57F first entered the rocket plume wakes ~15 min after launch at an altitude of ~18 km, then traversed the plume every ~5 min, climbing between plume passes to ensure that the same plume segment was not sampled twice. The WB-57F pilot identified plume segments for each pass and recorded aircraft ingress and egress times on the basis of visual identification of plume boundaries. The durations of the plume passes were highly variable, lasting from 5 to 60 s (900 m to 11 km at the WB-57F airspeed of ~180 m s⁻¹) and generally increased with time, reflecting expansion and diffusion of the plume.

Here we present data from selected plume traverses to illustrate the salient features of our observations. Figure 1 shows ozone concentrations measured during the first and eighth passes (out of a total of 18) through the 12 May Titan IV plume. The first pass occurred 16 min after launch at an altitude of 17.7 km; the observed ozone loss was ~35% across a 1.8-km plume traverse, similar to the ozone loss (40%) and plume width (1.5 km) reported 13 min after the 1975 Titan III launch¹¹. (The Titan III and Titan IV SRM exhaust composition and mass flow are similar.) Figure 2 shows ozone concentrations as measured at the time of the fifth and twelfth passes (out of a total of 13) through the 20 December Titan IV plume. The passes through this plume were generally shorter than those through the May 12 plume. Rawinsondes taken shortly before the launches show a more intense lower stratospheric wind shear environment on 20 December than on 12 May, indicating that turbulent diffusion on 20 December was more vigorous. Table 1 presents details of the plume wake dimension and ozone losses observed during the two measurement campaigns.

A general picture of the evolution of ozone concentration in a Titan IV plume wake can be drawn for the entire collection of plume data. During the first tens of minutes after launch, ozone loss reaches several tens of percent in narrow regions a few kilometres across (Fig. 1a and Table 1). During the following period (30–60 min after launch), the plume expands at a rate of about 0.1 km min⁻¹ and the most severe disturbances take place, with ozone losses approaching 100% over regions reaching 8 km across (Figs 1b and 2a; Table 1). After this time, as the plume continues to expand, the relatively large, deeply depleted regions are no longer detected and ozone concentrations in the plume begin to return to ambient levels (Fig. 2b and Table 1). This indicates that the ability of

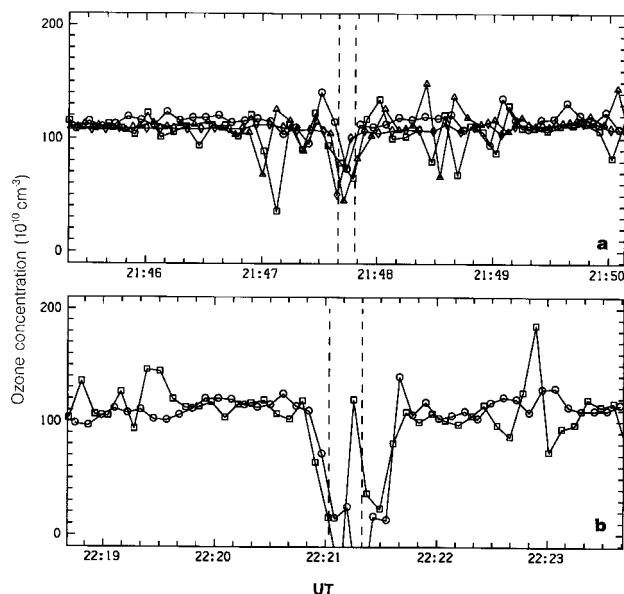


Figure 1 Measured ozone concentration during the time of WB-57F passes 1 (a) and 8 (b) through a Titan IV plume on 12 May 1996. Vertical lines indicate aircrew report of plume ingress and egress. Launch was at 21:33 UT. Circles and squares refer to channels 1 and 2 of the R instrument, respectively. Triangles and diamonds refer to channels 1 and 2 of the B instrument, respectively. Each instrument reports absolute ozone concentration from each of two channels every 8 s; staggered report times from both instruments provide a measurement every ~4 s. The r.m.s. deviations of the R1, R2, B1 and B2 channels around the time of, but not including, pass 1 are 7.5×10^{10} , 13.1×10^{10} , 14.7×10^{10} and 2.5×10^{10} cm⁻³, respectively. The r.m.s. deviations of the R1 and R2 channels around the time of, but not including, pass 8 are 8.3×10^{10} and 21×10^{10} cm⁻³, respectively. The B instrument failed before pass 8.

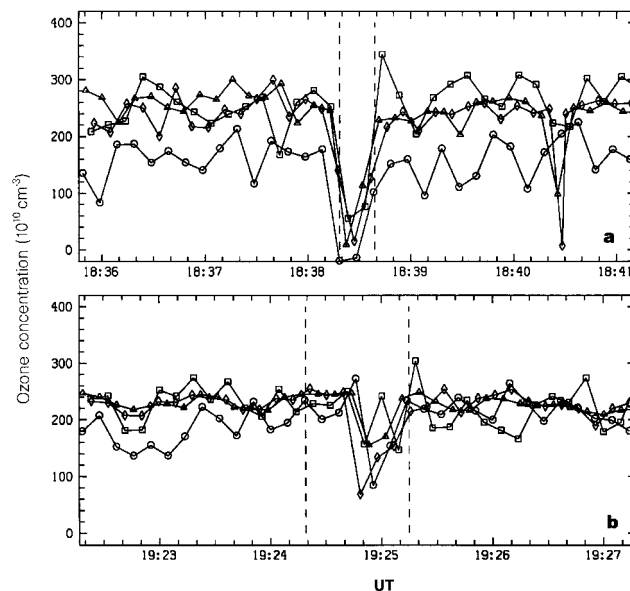


Figure 2 Measured ozone concentration during the time of WB-57F passes 5 (a) and 12 (b) through a Titan IV plume on 20 December 1996. Launch was at 18:04 UT. Symbols are as for Fig. 1. The r.m.s. deviations of the R1, R2, B1 and B2 channels around the time of, but not including, pass 5 are 37×10^{10} , 33×10^{10} , 38×10^{10} and 52×10^{10} cm⁻³, respectively. The r.m.s. deviations of the R1, R2, B1 and B2 channels around the time of, but not including, pass 12 are 34×10^{10} , 35×10^{10} , 10×10^{10} and 15×10^{10} cm⁻³, respectively.

plume gases to destroy ozone is spent 60 min after launch, and ozone-rich air can diffuse back into the plume wake to replace the lost ozone. Because ozone production is very slow in the lower stratosphere, ozone is replaced in the plume only in a relative sense. Still, the observed behaviour of ozone in the plume wake, in conjunction with the distortion of the plume from stratospheric wind shear, implies that the potential for significant changes in the total ozone column or solar ultraviolet exposure near launch sites is extremely limited. Accordingly, the local environmental hazard from transient stratospheric ozone loss after solid-fuelled rocket launches is not significant.

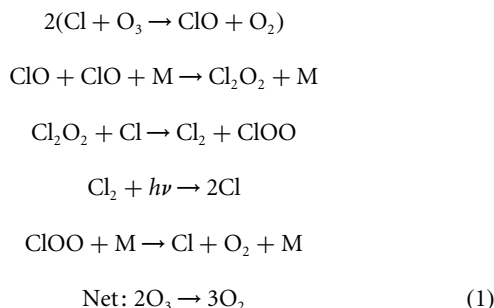
The Titan IV data seem to verify models that predict marked temporary ozone depletion in SRM plumes owing to reactive chlorine. However, the data do not provide direct evidence that chlorine is the main agent causing ozone loss. Strong support for a chlorine-based loss mechanism was obtained from night-time measurements in a Titan IV plume, which were carried out at stratospheric altitudes during a WB-57F flight on 24 April 1996. Twenty-nine minutes after launch, ozone loss was found to be negligible, despite Cl_2 concentrations reaching $\sim 15\%$ of ambient ozone levels¹⁶. The absence of significant ozone depletion after the night-time launch indicates that sunlight is required to drive the dominant ozone destruction reaction(s). This observation excludes exhaust NO or OH from being important ozone-destroying species in the plume. The presence of NO_2 , in contrast, would result in a photocatalytic ozone destruction cycle, but NO_2 has never been considered as a significant SRM exhaust component (ref. 2 and P. Zittel, unpublished data). These considerations, in conjunction with the direct observation of significant concentrations of Cl_2 (which is photolytically readily converted into reactive chlorine, Cl) in the plume, strongly implicate chlorine as causing the severe transient ozone losses in the daytime plume wakes.

Models predict that the Titan IV deposits ~ 5.5 mol of Cl_2 per metre of altitude in the lower stratosphere, representing $\sim 30\%$ of all chlorine in the exhaust¹. Detailed analysis of Fig. 1b reveals that ~ 40 mol of ozone had been lost per metre of altitude by the time of this pass. (Because the vehicle trajectory is nearly vertical and stratospheric turbulent mixing is horizontally isotropic, we assume that the WB-57F crossed the plume along the diameter of a roughly circular plume region). Data from the 20 December plume present a similar situation: 34 min after launch the plume was only 3.8 km across (Fig. 2a), but the total number of ozone molecules lost was similar to that in the 12 May plume because the ambient ozone concentration was greater than on 12 May. Thus, lacking other ozone-destroying species in the exhaust, each chlorine atom must have destroyed about four ozone molecules in the first 35–50 min after launch. The apparent disparity between the lost ozone and injected reactive chlorine is all the more evident because not all of the exhaust Cl will be available for ozone destruction. The mixing ratio of methane in the lower stratosphere is typically comparable to that of ozone¹⁷, and reaction with methane will convert some exhaust chlorine into HCl, the principal stratospheric chlorine sink.

If 100% of the exhaust chlorine appeared as Cl_2 , instead of the assumed 30%, the greater part of the observed loss could be explained. But plume models do not predict 100% conversion of HCl into Cl_2 at 18 km altitude¹. Several other mechanisms might contribute to ozone removal in the Titan IV exhaust wakes: reaction with exhaust nitrogen radicals, unspecified heterogeneous reactions on the surface of the exhaust alumina, or a fast catalytic cycle involving chlorine. Models predict only very small NO_x production in the stratosphere, less than 1% exhaust mass fraction for the Titan IV (P. Zittel, unpublished data), not enough to contribute significantly to the ozone loss. The potential significance of heterogeneous reactions on alumina particulates is more difficult to assess. The size distribution and effective surface area are not well known and a complete description of all the relevant reactions is not available.

One possibility, ozone decomposition on alumina, has recently been shown to be not significant¹⁸.

A gas-phase catalytic cycle that could possibly explain the magnitude of the measured ozone destruction involves the ClO dimer^{4,7}. The proposed cycle in the plume is given by:



Shortly after SRM plume deposition, when ClO and Cl concentrations are predicted to be of order one hundred parts per billion volume (p.p.b.v.)^{5,7}, the reaction of Cl_2O_2 with Cl will dominate over competing Cl_2O_2 loss mechanisms, including photolysis and thermal decomposition (despite the relatively warm ambient temperature of ~ 220 K, which promotes thermal decomposition)¹⁹. The cycle is expected to be rate-limited by Cl_2 photolysis proceeding at a rate of 0.13 min^{-1} , fast enough to allow each Cl atom to destroy several ozone molecules during the ~ 30 -min duration of intense ozone removal observed in the Titan IV plumes and consistent with the estimated ratio of ozone lost to chlorine injected, ~ 4 . Without additional information on plume composition we cannot unequivocally ascribe a major part of the observed ozone removal to this catalytic cycle; resolution of the problem awaits more detailed measurements *in situ*. □

Received 16 December; accepted 9 September 1997.

1. Zittel, P. F. *Computer Model Predictions of the Local Effects of Large Solid Fuel Rocket Motors on Stratospheric Ozone* (TR-94(4231)-9, The Aerospace Corporation, El Segundo, CA, 1994).
2. Karol, I. L., Ozolin, Y. E. & Rozanov, E. V. Effect of space rocket launches on ozone and other atmospheric gases. *Ann. Geophys.* **10**, 810–814 (1992).
3. Danilin, M. Y. Local stratospheric effects of solid-fueled rocket emission. *Ann. Geophys.* **11**, 828–836 (1993).
4. Kruger, B. C. Ozone depletion in the plume of a solid rocket motor. *Ann. Geophys.* **12**, 409–416 (1994).
5. Martin, L. R. *Possible Effect of the Chlorine Oxide Dimer on Transient Ozone Loss in Rocket Plumes* (RT-94(4231)-1, The Aerospace Corporation, El Segundo, CA, 1994).
6. Denison, M. R., Lamb, J. J., Bjorndahl, W. D., Wong, E. Y. & Lohn, P. D. Solid rocket motor exhaust in the stratosphere. *J. Spacecraft Rockets* **31**, 435–442 (1994).
7. Ross, M. N. Local effects of solid rocket motor exhaust on stratospheric ozone. *J. Spacecraft Rockets* **33**, 144–153 (1996).
8. Ko, M. K., Sze, N. D. & Prather, M. J. Better protection for the ozone layer. *Nature* **367**, 505–509 (1994).
9. *Scientific Assessment of Ozone Depletion*, Ch. 10 (Rep. no. 25, World Meteorological Organization, Geneva, Switzerland, 1991).
10. Jackman, C. H., Considine, D. B. & Fleming, E. L. The Space Shuttle's impact on the stratosphere: an update. *J. Geophys. Res.* **101**, 12523–12529 (1996).
11. Pergament, H. S., Gombert, R. I. & Poppoff, I. G. *Proceedings of the Space Shuttle Environmental Assessment Workshop on Stratospheric Effects* (Tech. Memo X-58198, NASA, Houston, TX, 1977).
12. McPeters, R., Prather, M. J. & Doiron, S. Reply to 'Comment on the Space Shuttle's impact on the stratosphere'. *J. Geophys. Res.* **96**, 17379–17381 (1991).
13. Syage, J. A. & Ross, M. N. An assessment of the TOMS for measuring ozone levels in a solid rocket motor plume. *Geophys. Res. Lett.* **23**, 3227–3230 (1996).
14. Sen, B., Benbrook, J. R. & Sheldon, W. R. Ultraviolet absorption photometer for measurement of ozone on a rocket-boostered payload. *Appl. Optics* **35**, 610–614 (1996).
15. Sheldon, W. R., Benbrook, J. R. & Amedieu, P. Ozone depletion in the upper stratosphere at the dawn terminator. *J. Atmos. Solar Terr. Phys.* **59**, 1–7 (1997).
16. Ross, M. N. *et al.* In-situ measurement of Cl_2 and O_3 in a stratospheric solid rocket motor exhaust plume. *Geophys. Res. Lett.* **24**, 1755–1758 (1997).
17. Roche, A. E. *et al.* Validation of CH_4 and N_2O measurements by the CLAES instrument on the Upper Atmospheric Research Satellite. *J. Geophys. Res.* **101**, 9679–9791 (1996).
18. Hanning-Lee, M. A., Brady, B. B., Martin, L. R. & Syage, J. A. Ozone decomposition on alumina: implications for solid rocket motor exhaust. *Geophys. Res. Lett.* **23**, 1961–1965 (1996).
19. Molina, L. T. & Molina, M. J. Production of ClO_2 from self-reaction of the ClO radical. *J. Phys. Chem.* **91**, 433–436 (1987).

Acknowledgements. We thank the NASA JSC WB-57F air and engineering crews, the ARES Mission Planning Office, and range support personnel at Vandenberg Air Force Base and Dryden Flight Research Center, for their efforts. This work is supported by the Launch Programs Office and the Office of Scientific Research of the United States Air Force.

Correspondence and requests for materials should be addressed to M.N.R. (e-mail: ross@courier3.aero.org).

Organic carbon burial forcing of the carbon cycle from Himalayan erosion

Christian France-Lanord*† & Louis A. Derry†

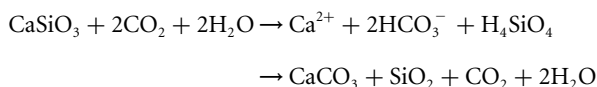
* Centre de Recherches Pétrographiques et Géochimiques, CNRS, BP20 54501 Vandoeuvre-les-Nancy, France

† Cornell University, Department of Geological Sciences, Ithaca, New York 14853, USA

Weathering and erosion can affect the long-term ocean–atmosphere budget of carbon dioxide both through the consumption of carbonic acid during silicate weathering and through changes in the weathering and burial rates of organic carbon^{1–4}. Recent attention has focused on increased silicate weathering of tectonically uplifted areas in the India–Asia collision zone as a possible cause for falling atmospheric CO₂ levels in the Cenozoic era^{5–7}. The chemistry of Neogene sediments from the main locus of sedimentary deposition for Himalayan detritus, the Bengal Fan, can be used to estimate the sinks of CO₂ from silicate weathering and from the weathering and burial of organic carbon resulting from Himalayan uplift. Here we show that Neogene CO₂ consumption from the net burial of organic carbon during Himalayan sediment deposition was 2–3 times that resulting from the weathering of Himalayan silicates. Thus the dominant effect of Neogene Himalayan erosion on the carbon cycle is an increase in the amount of organic carbon in the sedimentary reservoir, not an increase in silicate weathering fluxes.

Silicate weathering is typically incongruent, yielding both a solute and a secondary mineral phase, so direct evidence of chemical weathering can be found in the record of secondary minerals in sedimentary basins. The Bengal Fan and Ganges–Brahmaputra (GB) delta contain a huge volume of sediment derived from erosion of the India–Asia collision zone, with 6×10^6 km³ deposited in the past 20 Myr (ref. 8). Isotopic data for Nd, Sr and O from Bengal Fan sediments show that the source for over 80% of the detritus since 20 Myr ago has been the high-grade metasedimentary rocks of the High Himalayan crystalline (HHC) sequence⁹. Clastic and carbonate sediments of the Precambrian Lesser Himalaya (LH) and Palaeozoic–Mesozoic Tethyan Himalaya (TH) are the other important sources of sediment to the Bengal Fan during the Neogene.

Carbon dioxide consumption from silicate weathering can be represented schematically by:



where 1 mol of CO₂ is sequestered as marine carbonate for each mol Ca (or Mg) derived from silicate dissolution. Weathering of Na and K silicates contributes a smaller fraction to the weathering CO₂ sink because alkalis can exchange for Ca ± Mg adsorbed on detrital clays in estuarine zones, and during alteration of the oceanic crust^{10,11}. We estimate the CO₂ consumption from silicate weathering by comparing the chemistry of the weathered sediments deposited in the Bengal Fan with the chemistry of their unaltered Himalayan source rocks. The comparison slightly overestimates CO₂ consumption because any base cations released by weathering with strong acids (such as H₂SO₄) are still included in this CO₂ consumption budget. To represent the source of Bengal Fan sediment we use a composite of 99 samples from outcrops in the HHC (Table 1). Adding samples from the LH and TH strata to the average value for Himalayan source rocks does not change the estimated weathering fluxes significantly, because the combined contribution from these units

to the sediment flux in the Bengal Fan is <20%, and all three (meta)sedimentary units are chemically similar. We analysed a subset of HHC samples from Central Nepal for total organic carbon (C_{org}) contents. Metamorphic rocks of the HHC average $0.05 \pm 0.03\%$ C_{org}. Sediments of the Lesser Himalaya also have low C_{org} contents, $\leq 0.10\%$, except for rare black shale beds¹². Sediments of the Tethyan Himalaya include both carbonates and Palaeozoic clastic sediments with low C_{org} values, and some Mesozoic shales with up to 1.5% C_{org} (ref. 13). We estimate a volume-weighted mean C_{org} content of $0.10 \pm 0.05\%$ for the source rocks of Bengal Fan sediment.

We sampled Himalayan-derived sediments from late Pleistocene to mid-Miocene age recovered from the distal Bengal Fan on Leg 116 of the Ocean Drilling Program¹⁴. The sediments were chosen to represent a range of weathering intensities based on clay mineralogy¹⁵. Before 7 Myr and after 1 Myr ago the clay mineral assemblage (<2 μm size fraction) in the Bengal Fan is dominantly illite plus chlorite (the IC assemblage). From 7 to 1 Myr, clays in the Fan are dominantly pedogenic smectite and kaolinite (SK assemblage), reflecting more intense weathering in the GB floodplain¹⁶. The fine-grained, SK sediments from the Bengal Fan are the most intensely weathered Himalayan sediments, and are less abundant than the IC sediments. Before dissolution, whole rock samples were disaggregated in distilled H₂O, then leached with 10% acetic acid to remove minor carbonates (diagenetic, biogenic and detrital). Leaching of diagenetic carbonate removes some Ca and Mg derived from alteration of detrital silicates, although Sr isotopic data on the carbonate fraction show that most cations are derived from sea water. Ion exchange of H⁺ for adsorbed cations on clays may also release some silicate-derived cations. Thus our technique probably causes us to overestimate the silicate-derived alkalinity flux.

We normalized the major element concentrations of HHC source material and Bengal Fan sediment to their Al₂O₃ contents, on the assumption that during low-intensity weathering characteristic of the Himalayan drainage, aluminium is conservative. The global mean transport of dissolved Al is estimated to be <0.1% of suspended load transport¹⁷, making this a reliable assumption. The differences in ratios of major cations to Al₂O₃ between the HHC and Bengal Fan sediments can result from weathering losses,

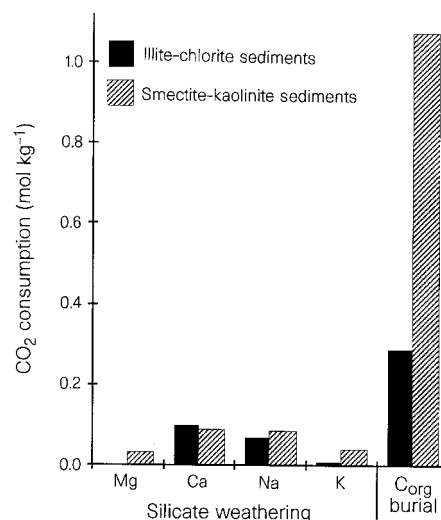


Figure 1 Consumption of atmospheric CO₂ by silicate weathering and burial of organic carbon calculated from the difference between the chemistry of the sediments deposited in the Bengal Fan and their unaltered Himalayan source rocks. Over the Neogene, the effect of silicate weathering is 2–3 times lower than the increase in the amount of C_{org} in the sedimentary reservoir.

Table 1 Average chemical compositions of the High Himalaya Crystalline (HHC) and decarbonated Bengal Fan sediments

Formation		SiO ₂	Al ₂ O ₃	Fe ₂ O ₃	MnO	MgO	CaO	Na ₂ O	K ₂ O	TiO ₂	P ₂ O ₅	LOI	Total	C _{org}
HHC <i>n</i> = 99	mean	71.87	13.11	4.36	0.07	1.76	1.07	2.26	3.03	0.62	0.04	1.26	99.43	0.05
	s.d.	6.62	2.97	1.92	0.03	1.19	0.81	0.90	1.06	0.28	0.08	1.01		0.03
Bengal IC <i>n</i> = 17	mean	62.82	16.71	6.42	0.04	2.23	0.68	1.14	3.64	0.72	0.14	5.11	99.60	0.44
	s.d.	8.42	3.73	2.15	0.01	0.62	0.50	0.51	0.90	0.16	0.04	2.12		0.50
Bengal SK <i>n</i> = 11	mean	54.10	17.83	9.73	0.05	2.23	0.81	0.75	2.19	1.28	0.15	10.63	99.76	1.39
	s.d.	5.55	1.50	2.16	0.04	0.38	0.46	0.41	0.92	0.54	0.03	3.94		0.62
Molar ratio		Si/Al	Fe/Al	Mn/Al	Mg/Al	Ca/Al	Na/Al	K/Al	Ti/Al					
HHC		2.33	0.21	0.00	0.08	0.04	0.28	0.25	0.02					
Bengal IC		1.59	0.25	0.00	0.08	0.02	0.11	0.24	0.01					
Bengal SK		1.29	0.35	0.00	0.08	0.02	0.07	0.13	0.02					

Major element analyses were made by inductively coupled plasma optical emission spectroscopy on samples fused with lithium metaborate.

C_{org}: HHC, *n* = 21; Bengal, *n* = 155 (refs 14, 20).

mineral sorting or residual enrichment of insoluble cations. Dissolved silica transport in the modern Ganges (from all sources) is about 0.3% of suspended load transport¹⁸, indicating that dissolution of quartz is minor. Low SiO₂/Al₂O₃ in the sediments results from mineral sorting, as quartz is mechanically resistant and is transported less readily than finer-grained aluminosilicate minerals. If fine-grained silicates are preferentially transported to the distal Bengal Fan our estimate of alkalinity flux is too high, as the fine-grained material contains a higher proportion of weathered clay minerals. However, the fine-grained material is also enriched in Mg, which tends to counteract this bias. Differences in the weathering intensity between the IC and SK sediments are apparent in the base cation/Al₂O₃ ratios. IC sediments have lost little or no K₂O and MgO relative to the HHC source rocks, and about half of their Na₂O and CaO. These results are in good agreement with analyses of modern soil profiles developed on Himalayan gneisses¹⁹. SK sediments have lost some MgO, about half of their K₂O and CaO, and much of their Na₂O. The mean C_{org} content of Bengal Fan sediments is 0.85% (*n* = 155), with the mean for IC sediments being 0.4%, whereas the mean for SK sediments is 1.5%²⁰. Data from fan sediments in the region near the delta yield similar values to the distal fan, with mean C_{org} = 0.90% (ref. 21), suggesting that any sorting effect on C_{org} contents is minor.

Consumption of CO₂ due to Himalayan silicate weathering was calculated assuming that all Mg²⁺ and Ca²⁺ lost from the silicates forms marine carbonates. The fraction of K and Na involved in cation exchange reactions with sediments or the oceanic crust remains poorly known^{10,22}. We conservatively estimate that 20% of K⁺ and 30% of Na⁺ in the global river flux exchanges for Ca ± Mg and produces carbonates. Accounting for charge balance, CO₂ consumption can be estimated from the base cation losses during Himalayan erosion:

$$-\Delta\text{CO}_2 = \Delta\text{Ca} + \Delta\text{Mg} + 0.10\Delta\text{K} + 0.15\Delta\text{Na}$$

We have ignored neoformation of marine clays ('reverse weathering'), which may be a significant sink for base cations²³ and would lower the estimated CO₂ consumption. Consumption of CO₂ by silicate weathering is 0.17 mol kg⁻¹ for the less weathered IC sediments and 0.23 mol kg⁻¹ for the SK sediments. Similarly, we estimate the CO₂ consumption from organic carbon (C_{org}) burial by comparing the C_{org} content of the clastic sediments in the Bengal Fan with the C_{org} content of their Himalayan source rocks. For IC sediments, this sink is about 0.27 mol CO₂ kg⁻¹; for SK sediments the sink is 1.1 mol CO₂ kg⁻¹, 1.7 and 4.7 times their respective silicate weathering sinks (Fig. 1). The more intensely weathered SK sediments also have higher C_{org} contents, more than offsetting the greater cation losses. These averages are representative of the range of sediment organic and inorganic chemistries in the Bengal Fan. We emphasize that this conclusion is based on a probable overestimate of the magnitude of the silicate weathering flux for CO₂:

our sample preparation and data analysis tends to bias this estimate towards high values. The result is independent of the magnitude of sediment fluxes past or present, because we have measured CO₂ consumption by inorganic and organic sinks in the same samples.

The fluxes of CO₂ implied by these estimates can be compared with independent global estimates. For a mean suspended load flux of 8×10^{11} kg yr⁻¹ (ref. 24) our data yield an average CO₂ consumption by silicate weathering in the GB system of 0.17×10^{12} mol yr⁻¹ during the Neogene. This flux is 2.6% of the current global silicate weathering sink for CO₂ (ref. 2). The modern GB contributes 2.7% of the global river water discharge and 2.1% of the global riverine SiO₂ flux^{18,22}. Thus our estimates for both the long-term discharge-weighted CO₂ consumption via silicate weathering and the modern discharge-weighted SiO₂ flux in the GB system are near the current world average. Despite the huge erosional flux from the Himalaya, the silicate weathering sink for CO₂ is modest on the global scale. The extreme relief of the Himalaya and the monsoon climate result in very rapid physical denudation and fast transport of sediment to the ocean. One result is a strongly weathering-limited system in which the kinetics of chemical weathering are slow relative to the transport time of eroded rock to the sea. Furthermore, Ca silicates are not abundant in the Himalaya, and the silicate-derived alkalinity flux is largely in the form of Na and K cations which are inefficient sinks of CO₂.

The data above yield an average Neogene rate of net growth (burial – weathering > 0) of the sedimentary C_{org} reservoir of 0.58×10^{12} mol yr⁻¹ in the Himalayan–Bengal system. The net growth in the size of the global sedimentary C_{org} reservoir can be estimated from the marine carbon isotopic mass balance²⁵. Our recent results from a δ¹³C model²⁶ yield a global average net flux to the sedimentary C_{org} reservoir of about 1.1×10^{12} mol yr over the past 15 Myr. The results are not directly comparable because one value represents a regional flux, whereas the other represents a global flux. However, they are consistent, and suggest that C_{org} burial in excess of weathering in the Himalayan–Bengal system can contribute significantly to changes in the global C_{org} reservoir. Together with the Indus fan and the Indo-Gangetic plain, the Bengal Fan accounts for about 15% of the modern total burial flux of global C_{org} (ref. 20), so it is not surprising that any imbalance (burial – weathering ≠ 0) in the Himalayan–Bengal C_{org} budget could have had a global impact. Up to 90% of C_{org} burial takes place in continental margin sediments²⁷, so any process that increases continental margin sedimentation significantly, such as erosion of a major orogen, may be expected to increase C_{org} burial and possibly amplify imbalances in the C_{org} budget. Rapid erosion and the high suspended load of the GB system help drive C_{org} burial rates high enough to perturb the global carbon cycle significantly. Erosion of a major orogenic belt such as the Himalaya creates a large amount of mineral surface area, which is a strong control on organic carbon burial in continental margin settings²⁸.

Recent work on the evolution of the global climate during the Cenozoic era has focused almost exclusively on the possible perturbation of atmospheric CO₂ levels resulting from weathering of silicates, especially in the Himalaya^{5,6,29,30}. But Himalayan erosion produces very large C_{org} fluxes^{20,31,32}. Although the hypothesized link between Himalayan silicate weathering and atmospheric CO₂ levels remains poorly quantified, our results indicate that increased sedimentary C_{org} storage resulting from Neogene Himalayan erosion and weathering has had a significantly larger effect on the carbon cycle than silicate weathering, by a factor of 2–3. Both models of the net change in the global sedimentary C_{org} reservoir^{7,26} and data from the Himalayan–Bengal system are consistent with the hypothesis that an excess of C_{org} burial over weathering acted as a sink for atmospheric CO₂ during the Neogene. □

Received 13 February; accepted 18 August 1997.

- Garrels, R. M. & Perry, E. A. Cycling of carbon sulfur and oxygen through geologic time. in *The Sea* 303–336 (1974).
- Lasaga, A. C., Berner, R. A. & Garrels, R. M. in *The Carbon Cycle and Atmospheric CO₂: Natural Variations, Archean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 397–410 (American Geophysical Union, Washington DC, 1985).
- Berner, R. A. Atmospheric CO₂ levels over Phanerozoic time. *Science* **249**, 1382–1386 (1990).
- Walker, J. C. G., Hays, P. B. & Kasting, J. F. A negative feedback mechanism for the long-term stabilization of Earth's surface temperature. *J. Geophys. Res.* **86**, 9776–9782 (1981).
- Raymo, M. E. & Ruddiman, W. F. Tectonic forcing of late Cenozoic climate. *Nature* **359**, 117–122 (1992).
- Molnar, P., England, P. & Martinod, J. Mantle dynamics, uplift of the Tibetan Plateau, and the Indian monsoon. *Rev. Geophys.* **31**, 327–396 (1993).
- Goddéris, Y. & François, L. M. Balancing the Cenozoic carbon and alkalinity cycles: Constraints from isotopic records. *Geophys. Res. Lett.* **23**, 3743–3746 (1996).
- Einsele, G., Ratschbacher, L. & Wetzel, A. The Himalaya–Bengal fan denudation accumulation system during the past 20 Ma. *J. Geol.* **104**, 163–184 (1996).
- France-Lanord, C., Derry, L. & Michard, A. in *Himalayan Tectonics* (eds Treloar, P. J. & Searle, M.) 603–621 (Geol. Soc. Lond., London, 1993).
- Sayles, F. L. & Mangelsdorf, P. C. The equilibration of clay minerals with seawater: exchange reactions. *Geochim. Cosmochim. Acta* **41**, 951–960 (1977).
- Von Damm, K. L., Edmond, J. M., Grant, B. & Measures, C. K. Chemistry of submarine hydrothermal solutions at 21°N, East Pacific Rise. *Geochim. Cosmochim. Acta* **49**, 2197–2220 (1985).
- France-Lanord, C. Chevauchement, métamorphisme et magmatisme en Himalaya du Népal central. Etude isotopique H, C, O. Thesis, Institut National Polytechnique de Lorraine (1987).
- Beck, R. A., Burbank, D. W., Sercombe, W. J., Olson, T. L. & Khan, A. M. Organic carbon exhumation and global warming during the early Himalayan collision. *Geology* **23**, 387–390 (1995).
- Shipboard Scientific Party. Initial Reports Sites 717–718–719 Distal Bengal Fan, *Proc. ODP Init. Rep.* (Ocean Drilling Program, College Station, TX, 1989).
- Bouquillon, A., France-Lanord, C., Michard, A. & Tiercelin, J.-J. in *Proc. ODP Sci. Res.* (eds Cochran, J. R. *et al.*) 43–58 (1990).
- Derry, L. A. & France-Lanord, C. Neogene Himalayan weathering history and river ⁸⁷Sr/⁸⁶Sr: Impact on the marine Sr record. *Earth Planet. Sci. Lett.* **142**, 59–74 (1996).
- Martin, J. M. & Meybeck, M. Element mass-balance of material carried by major world rivers. *Mar. Chem.* **7**, 173–206 (1979).
- Sarin, M. M., Krishnaswami, S., Dilli, K., Omayajulu, B. I. K. & Moore, W. S. Major ion chemistry of the Ganga–Brahmaputra river system: Weathering processes and fluxes to the Bay of Bengal. *Geochim. Cosmochim. Acta* **53**, 997–1009 (1989).
- Gardner, R. & Walsh, N. Chemical weathering of metamorphic rocks from low elevations in the southern Himalaya. *Chem. Geol.* **127**, 161–176 (1996).
- France-Lanord, C. & Derry, L. A. δ¹³C or organic carbon in the Bengal fan: source evolution and transport of C3 and C4 plant carbon to marine sediments. *Geochim. Cosmochim. Acta* **58**, 4809–4814 (1994).
- Bouquillon, A. *Influence continentale et marine dans les sédiments Cénozoïques de l'océan Indien Nord Oriental*. Thesis, Université de Lille Flandres-Artois (1987).
- Berner, R. A. & Berner, E. *Global Environment: Water, Air, and Geochemical Cycles* 1–376 (Prentice Hall, Upper Saddle River, NJ, 1994).
- Michalopoulos, P. & Aller, R. C. Rapid clay mineral formation in Amazon delta sediments: Reverse weathering and oceanic elemental cycles. *Science* **270**, 614–617 (1995).
- Milliman, J. D. & Syvitski, P. M. Geomorphic/tectonic control of sediment discharge of the ocean: the importance of small mountainous rivers. *J. Geol.* **100**, 525–544 (1992).
- Shackleton, S. J. & Hall, M. A. Carbon isotope data from Leg 74 sediments. in *Init. Rep. DSDP* (eds Moore, T. C. *et al.*) 613–619 (US Govt Printing Office, Washington, 1984).
- Derry, L. A. & France-Lanord, C. Neogene growth of the sedimentary organic carbon reservoir. *Paleoceanography* **11**, 267–275 (1996).
- Berner, R. A. & Canfield, D. A model for atmospheric oxygen over Phanerozoic time. *Am. J. Sci.* **289**, 333–361 (1989).
- Hedges, J. I. & Keil, R. G. Sedimentary organic matter preservation: an assessment and speculative synthesis. *Mar. Chem.* **49**, 81–90 (1995).
- Raymo, M. E. The Himalayas, organic carbon burial, and climate in the Miocene. *Paleoceanography* **9**, 399–404 (1994).
- Caldeira, K. Enhanced Cenozoic chemical weathering and the subduction of pelagic carbonate. *Nature* **357**, 578–581 (1992).
- Subramanian, V. & Ittekkot, V. Carbon transport by the Himalayan rivers. in *Biogeochemistry of Major World Rivers, SCOPE* (eds Degens, E. T., Kempe, S. & Richey, J. E.) 157–168 (Wiley, Chichester, 1991).
- Ittekkot, V. Global trends in the nature of organic matter in river suspensions. *Nature* **332**, 436–438 (1988).

Acknowledgements. We thank Patrick Le Fort for providing part of the Himalayan samples and analyses. This study was supported by the CNRS-INSU program 'Dynamique et Bilan de la Terre-Fluve et Érosion'.

Correspondence should be addressed to C.F.-L. at CRPG-CNRS (e-mail: cfl@crpg.cnrs-nancy.fr).

Mechanical power output of bird flight

K. P. Dial*, A. A. Biewener†, B. W. Tobalske* & D. R. Warrick*

* Division of Biological Sciences, University of Montana, Missoula, Montana 59812, USA

† Department of Organismal Biology and Anatomy, University of Chicago, 1025 East 57 Street, Chicago, Illinois 60637, USA

Aerodynamic theory predicts that the power required for an animal to fly over a range of speeds is represented by a 'U'-shaped curve, with the greatest power required at the slowest and fastest speeds, and minimum power at an intermediate speed^{1–6}. Tests of these predictions, based on oxygen consumption measurements of metabolic power in birds^{7–12} and insects¹³, support a different interpretation, generating either flat or 'J'-shaped power profiles, implying little additional demand between hovering and intermediate flight speeds¹⁴. However, respirometric techniques represent only an indirect assessment of the mechanical power requirements of flight and no previous avian study has investigated an animal's full range of attainable level flight speeds. Here we present data from *in vivo* bone-strain measurements of

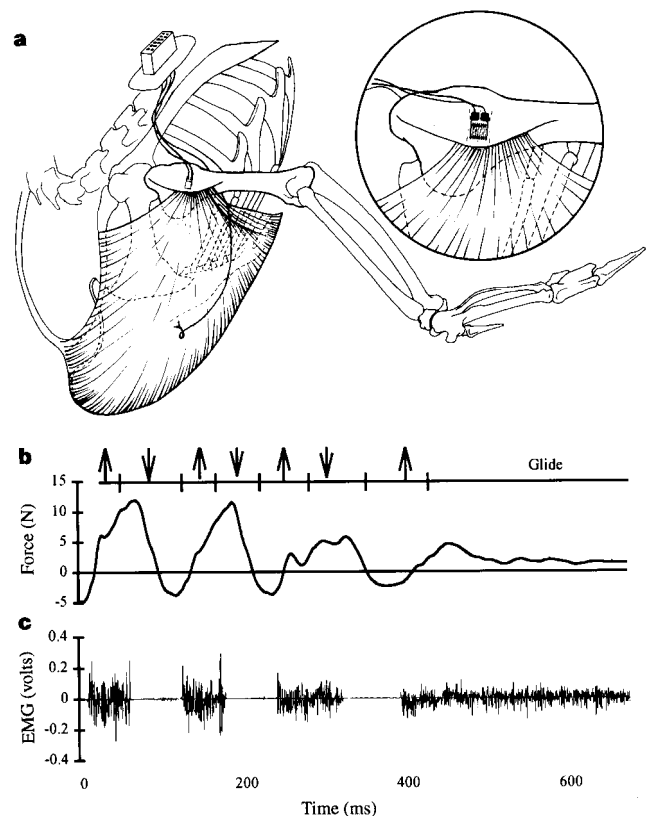


Figure 1 **a**, Illustration of a magpie pectoral girdle and forelimb showing the placement of a strain gauge on the delto-pectoral crest (DPC; expanded in inset) and indwelling bipolar electromyographic (EMG) wires in the pectoralis muscle, all connected to a dorsal plug. **b**, Recordings of DPC strain calibrated to pectoralis force (arrows indicate kinematic upstroke and downstroke) and **c**, corresponding pectoralis (EMG) for three successive wingbeats of a magpie flying at 6 m s⁻¹, were synchronized to wing kinematics (obtained from 16 mm high-speed movie film at 150–200 f.p.s. in lateral and dorsal views) to estimate pectoralis fibre length change in relation to force development and to determine wingbeat cycle duration.

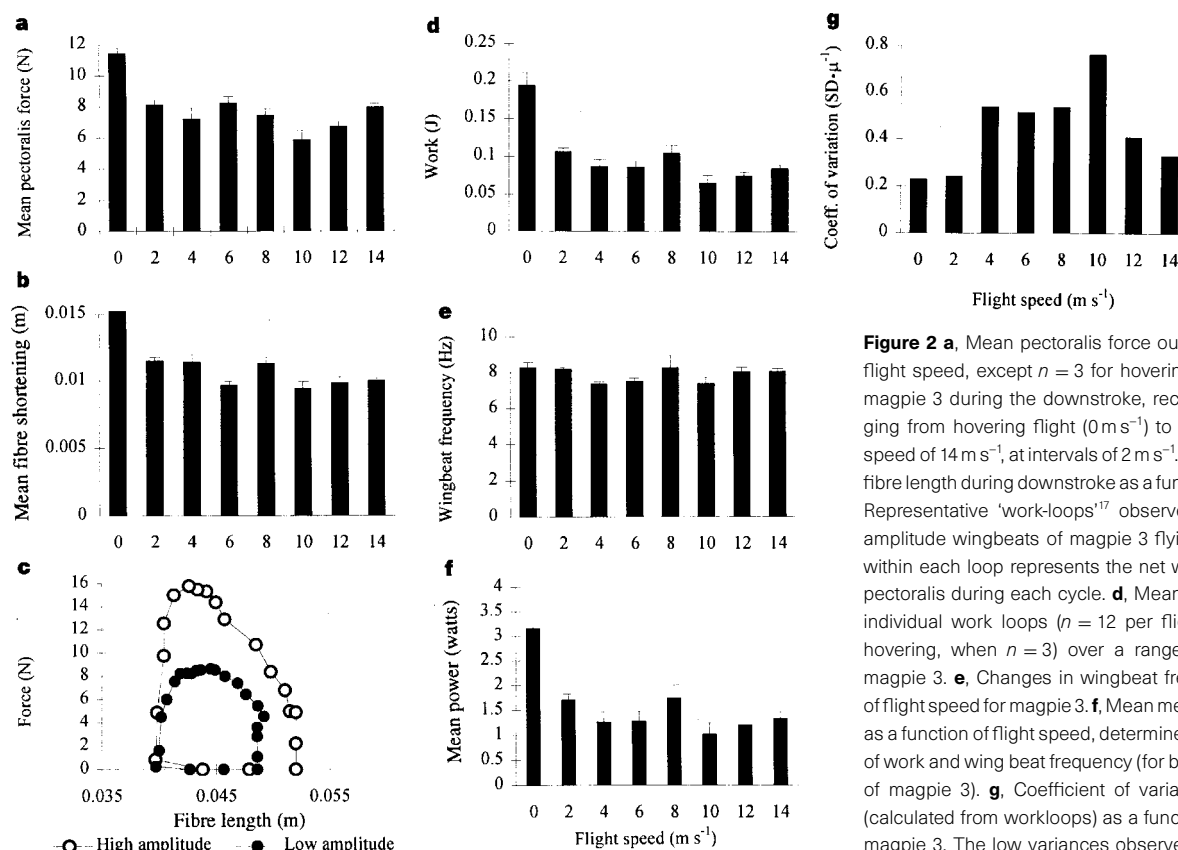


Figure 2 **a**, Mean pectoralis force output ($\pm$ s.e.; $n = 12$ per flight speed, except $n = 3$ for hovering flight) generated by magpie 3 during the downstroke, recorded for speeds ranging from hovering flight (0 m s^{-1}) to a maximum observed speed of 14 m s^{-1} , at intervals of 2 m s^{-1} . **b**, Changes in muscle-fibre length during downstroke as a function of flight speed. **c**, Representative 'work-loops'¹⁷ observed for high- and low-amplitude wingbeats of magpie 3 flying at 6 m s^{-1} ; the area within each loop represents the net work performed by the pectoralis during each cycle. **d**, Mean work calculated from individual work loops ($n = 12$ per flight speed, except for hovering, when $n = 3$) over a range of flight speeds for magpie 3. **e**, Changes in wingbeat frequency as a function of flight speed for magpie 3. **f**, Mean mechanical power output as a function of flight speed, determined from measurements of work and wing beat frequency (for both pectoralis muscles of magpie 3). **g**, Coefficient of variation for power output (calculated from workloops) as a function of flight speed for magpie 3. The low variances observed at hovering and the highest flight speeds suggest that the bird was performing at or near to its performance limit. Comparable data to those shown here were obtained for the other two magpies.

pectoralis muscle force coupled with wing kinematics in black-billed magpies (*Pica pica*), which we use to calculate mechanical power directly. As these birds flew over their full range of speeds, we offer a complete profile of mechanical power output during level flapping flight for this species. Values of mechanical power output are statistically indistinguishable (that is, the power curve is flat) over most forward-flight speeds but are significantly higher during hovering and flight at very low speeds.

We flew birds in a variable-speed wind tunnel^{15,16} to measure the force generated by the dominant flight muscles by means of bone-strain recordings (Figs 1, 2a) and to determine muscle-fibre shortening from analysis of wing excursion (Fig. 2b). Changes in force plotted against changes in fibre length (Fig. 2c) show the net total work (area of 'work loop')¹⁷ of the pectoralis during a wingbeat cycle. By combining measurements of work per cycle (Fig. 2d) with wingbeat frequency (Fig. 2e), we calculated the mechanical power generated by the pectoralis muscles over a range of steady flight speeds. The results obtained from our measurements and analyses show that the mechanical power output (mean body-mass-specific power) is maximal during hovering ($\sim 21 \text{ W kg}^{-1}$ at 0 m s^{-1}), decreases markedly with an increase in forward speed ($\sim 9 \text{ W kg}^{-1}$ at 4 m s^{-1}), remains relatively constant over a broad range of faster flight speeds (4 to 12 m s^{-1}), and increases slightly during the maximum flight speed measured ($\sim 12 \text{ W kg}^{-1}$ at 14 m s^{-1}) (Figs 2, 3). This pattern is borne out by the generally uniform measurements of force, fibre shortening and wingbeat frequency (Fig. 2a, b, e) that we observed over the intermediate speed range, and the maximum pectoralis force generation and fibre shortening during hovering. Correspondingly, the low values in the coefficients of variation (CV) of power output at the lowest and highest flight speeds, and the more or less stereotypic wingbeat excursions, suggest that each animal was performing at or near its limit at

these speeds (Fig. 2g). In contrast, higher CV values at intermediate flight speeds reflect the use of variable wing excursions¹⁸, which resulted in constant power production over a wide range of flight speeds. We propose that alterations in wing and tail configuration during different flight speeds account for the differences in mechanical power otherwise predicted by quasi steady-state aerodynamic theory of vertebrate flight.

Our calculations of mechanical power output are consistent with estimates from previous metabolic studies^{8–12} at intermediate speeds. However, as previous metabolic data do not describe the energetic cost of flight over each species' full range of speeds (that is, from hovering to maximum flapping flight), the power profiles are incomplete (Fig. 3b). This is undoubtedly due to the difficulty (or impossibility) of obtaining reliable steady-state measurements of oxygen consumption during extremely slow flight speeds ($0\text{--}4 \text{ m s}^{-1}$). In contrast, the technique we describe provides instantaneous (that is, per wingbeat) measurements of muscle mechanical power during hovering and very slow flight.

Our measurements show that power output is maximal during hovering. Comparable levels of power output are not observed at the upper range of flying speeds. As hovering episodes are brief (1–4 s), we suggest that non-forward flapping flight in this species, as in other species, may involve anaerobic metabolism within the pectoralis. Similarly, sustainable flight at speeds greater than 14 m s^{-1} may not be possible because anaerobic metabolism cannot be sustained. Thus, the 'shape' of the power curve will be determined by the particular species' available aerobic power. Birds that can sustain higher levels of mass-specific power output may be more likely to fly at very slow and fast speeds, exhibiting high induced power (slow) or high profile and parasite power (fast)¹⁴.

We suggest that the 'L'-shaped power profile exhibited by magpies reflects the power required for birds of similar wing design and body

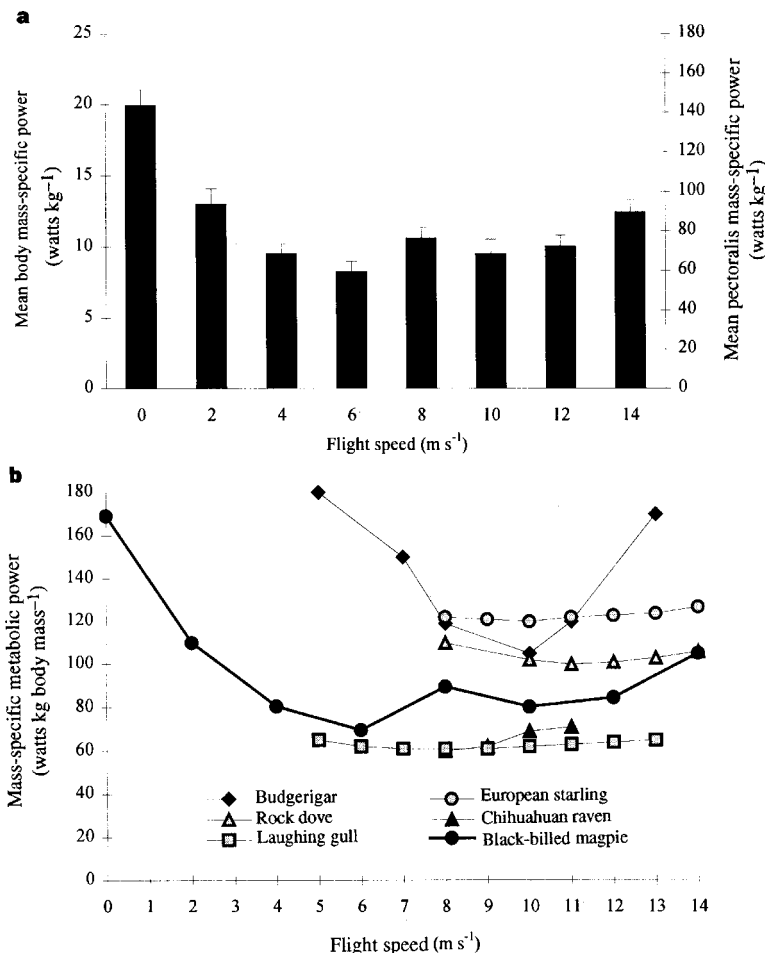


Figure 3 a, Summary of the mean total mechanical power output measured for all three magpies as a function of flight speed. Whereas power output during hovering was significantly higher ($P < 0.001$) than at all other speeds, no significant differences in power were observed between 4 and 14 m s⁻¹. Probability values were derived from repeated measures ANOVA with post-hoc comparison of power output versus speed ($n = 3$ or 10 wingbeats per flight speed). **b**, Comparison of our present data for mechanical power output of the magpie (bold line) with mass-specific metabolic power output data previously reported for various bird species⁶, based on a muscle efficiency of 11.75%. This value was obtained by averaging the results obtained for the metabolic^{11,12} and mechanical power^{15,18} output of starlings (13.0% at 11–14 m s⁻¹) and pigeons (10.5% at 7–9 m s⁻¹) during forward flight. All bird studies^{8–12}, with the exception of the budgerigar⁷, show power output during level flapping flight to be relatively flat (that is, constant) over most speeds recorded. Note that at low flight speeds (0 and 2 m s⁻¹), birds clearly recruit many additional muscles (such as tail and antebrachial muscles) that do no measurable mechanical work²⁵, so our technique may underestimate the metabolic power required at these speeds.

mass. However, the flight of other species of birds may not adhere to the 'U'-shaped power curve that we describe for magpies. Hummingbirds, for example, have a unique shoulder arrangement that permits a wing movement to generate lift during both 'upstroke' and 'downstroke', perhaps allowing members of this group to hover using significantly less mass-specific mechanical power output than other small birds^{5,14,19}. Testing the generality of the profile of mechanical power output for magpies requires that similar measurements be made for a number of avian species having disparate wing structure. The fact that birds do not appear to adhere to a strict 'U'-shaped power curve may help to explain why field observations^{20,21} frequently fail to agree with the theoretically predicted⁴ minimum power and maximum range flight velocities. Such theories do not adequately account for the ability of birds to dramatically change wing and tail presentation over a range of flight speeds¹⁸, permitting these animals to alter significantly the power requirements for flight. □

Methods

Animals and strain measurements. Following training to fly in a variable-speed wind tunnel^{16,18} over nearly their full range of flight speeds (hovering to 14 m s⁻¹), *in vivo* measurements of pectoralis muscle force were obtained from three black-billed magpies (mean mass, 174 g), based on bone-strain recordings made at the dorsal surface of the deltopectoral crest (DPC)^{15,22}. The DPC serves as the insertion site for the pectoralis on the proximo-anterior surface of the humerus and as such the DPC's dorsal surface experiences tensile strains as the pectoralis develops force. Our measurements assume that the power generated by the pectoralis muscles approximates the total mechanical power of the bird during flight^{23,24}. However, this method may slightly underestimate the mechanical power requirements during very slow or hovering flight as it ignores the inertial power generated by the supracoracoideus and/or other

upstroke muscles²⁵. The strain gauge was attached perpendicularly to the bone's shaft to avoid sensitivity to strains produced by bending and torsional loading of the humerus and associated with aerodynamic loading of the wing¹⁹. Following recovery from strain gauge and electrode implantation, birds were flown at speeds from 0 to 14 m s⁻¹ at 2 m s⁻¹ intervals. Birds were flown first at an intermediate speed, with subsequent speed trials conducted in a random order. DPC strains were then calibrated *in situ* in fully anaesthetized animals, using a force transducer attached to the pectoralis muscle by a heavy nylon thread tied around the muscle's aponeurotic insertion below the ventral surface of the DPC, so that cyclic forces could be applied in the direction of the muscle's pull during flight. Force calibrations were made with the wings held at 60°, 30° and horizontal in order to obtain a mean calibration. After calibration, the strain gauge and electromyogram electrodes were removed and the birds allowed to recover.

Measurements of muscle fibre length. Changes in muscle fibre length as a function of flight speed were estimated from measurements of wingbeat amplitude by digitizing wingtip elevation (h) and wing length from lateral and dorsal views of high-speed film to calculate the wing excursion angle, $\theta = \sin^{-1}(h/b)$, where b is the extended wing length at mid-downstroke. Pectoralis fibre length change (Δl) was then calculated as the arc of movement of the humerus at the DPC ($\Delta l = \Delta \theta r$), based on the measurements of wing excursion angle and the muscle's moment arm (r) at the shoulder. This method assumes that excursions of the wing and humerus are uniformly maintained with respect to one another over the animal's entire flight speed range. Relative changes in pectoralis fibre-length change were then calculated on the basis of the measurement of an average resting fibre length made directly from the pectoralis of one of the magpies used in the wind-tunnel trials.

Calculating work loops. Changes in force plotted against changes in fibre length provide the net total work of the pectoralis during a wingbeat cycle. Multiplying measurements of work per cycle by wingbeat frequency yields mechanical power generated by the muscles.

Received 11 June; accepted 11 August 1997.

1. Tucker, V. A. Bird metabolism during flight: evaluation of a theory. *J. Exp. Biol.* **58**, 689–709 (1973).
2. Pennycuik, C. J. Power requirements for horizontal flight in the pigeon *Columba livia*. *J. Exp. Biol.* **49**, 527–555 (1968).
3. Pennycuik, C. J. in *Avian Biology* (eds. Farner, D. S. & King, J. R.) **5**, 1–75 (1975).
4. Pennycuik, C. J. *Bird Flight Performance* (Univ. Press, Oxford, 1989).
5. Rayner, J. M. V. A new approach to animal flight mechanics. *J. Exp. Biol.* **80**, 17–54 (1979).
6. Norberg, U. M. *Vertebrate Flight* (Springer, Berlin, 1990).
7. Tucker, V. A. Respiratory exchange and evaporative water loss in the flying budgerigar. *J. Exp. Biol.* **48**, 67–87 (1968).
8. Tucker, V. A. Metabolism during flight in the laughing gull, *Larus atricilla*. *Am. J. Physiol.* **222**, 237–245 (1972).
9. Hudson, D. M. & Bernstein, M. H. Gas exchange and energy cost of flight in the white-necked raven, *Corvus cryptoleucus*. *J. Exp. Biol.* **103**, 121–130 (1983).
10. Berger, M., Hart, O. Z. & Roy, J. S. Respiration, oxygen consumption and heart rate in some birds during rest and flight. *Z. Vergl. Physiol.* **66**, 201–214 (1970).
11. Torre-Bueno, J. R. & LaRoche, J. The metabolic cost of flight in unrestrained birds. *J. Exp. Biol.* **75**, 223–229 (1978).
12. Rothe, H.-J., Biesel, W. & Nachtigall, W. Pigeon flight in a wind tunnel. II. Gas exchange and power requirements. *J. Comp. Physiol. B* **157**, 99–109 (1987).
13. Ellington, C. P., Machin, K. I. & Casey, T. M. Oxygen consumption of bumblebees in forward flight. *Nature* **347**, 472–473 (1990).
14. Ellington, C. P. Limitations on animal flight performance. *J. Exp. Biol.* **160**, 71–91 (1991).
15. Dial, K. P. & Biewener, A. A. Pectoralis muscle force and power output during different modes of flight in pigeons (*Columba livia*). *J. Exp. Biol.* **176**, 31–54 (1993).
16. Tobalske, B. W. & Dial, K. P. Neuromuscular control and kinematics of intermittent flight in budgerigars (*Melopsittacus undulatus*). *J. Exp. Biol.* **187**, 1–13 (1994).
17. Josephson, R. K. The mechanical power output of a tettigoniid wing muscle during singing and flight. *J. Exp. Biol.* **117**, 357–368 (1985).
18. Tobalske, B. W. & Dial, K. P. Flight kinematics of black-billed magpies and pigeons over a wide range of speeds. *J. Exp. Biol.* **199**, 263–280 (1996).
19. Biewener, A. A. & Dial, K. P. *In vivo* strain in the pigeon humerus during flight. *J. Morph.* **225**, 61–75 (1995).
20. Hedenstrom, A. & T. Ålerstam. Skylark optimal flight speeds for flying nowhere and somewhere. *Behav. Ecology* **7**, 121–126.
21. Liechti, F., Ehrlich, D. & Bruderer, B. Flight behaviour of white storks *Ciconia ciconia* on their migration over Southern Israel. *Ardea* **84**, 3–13 (1997).
22. Biewener, A. A., Dial, K. P. & Goslow, G. E. Jr Pectoralis muscle force and power output during flight in the starling. *J. Exp. Biol.* **164**, 1–18 (1992).
23. Dial, K. P. Activity patterns of the wing muscles in the pigeon (*Columba livia*) during different modes of flight. *J. Exp. Zool.* **262**, 357–373 (1992).
24. Dial, K. P. Avian forelimb muscles and nonsteady flight: can birds fly without using the muscles in their wings? *Auk* **109**, 874–885 (1992).
25. Van Den Berg, C. & Rayner, J. M. V. The moment of inertia of bird wings and the inertial power requirement for flapping flight. *J. exp. Biol.* **198**, 1655–1664 (1995).
26. Scholey, K. D. *Evolution in Vertebrate Flight: Climbing and Gliding of Mammals and Reptiles and the Flapping Flight of Birds*. thesis, Univ. Bristol (1983).

Acknowledgements. We thank F. A. Jenkins Jr for his thoughtful input on previous drafts of this manuscript; M. LaBarbera, J. M. Marzluff, D. Fawcett and D. F. Boggs for their comments; and J. Gilpin for making the force-transducer used to calibrate the DPC strain recordings. Supported by grants from the NSF to K.P.D. and A.A.B.

Correspondence and requests for materials should be addressed to K.P.D. (e-mail: kdial@selway.umd.edu).

Impaired odour discrimination on desynchronization of odour-encoding neural assemblies

Mark Stopfer*, Seetha Bhagavan†‡, Brian H. Smith† & Gilles Laurent*

* California Institute of Technology, Biology Division, 139-74, Pasadena, California 91125, USA

† Ohio State University, Department of Entomology, 1735 Neil Avenue, Columbus, Ohio 4210-1220, USA

Stimulus-evoked oscillatory synchronization of neural assemblies has been described in the olfactory^{1–5} and visual^{6–8} systems of several vertebrates and invertebrates. In locusts, information about odour identity is contained in the timing of action potentials in an oscillatory population response^{9–11}, suggesting that oscillations may reflect a common reference for messages encoded in time. Although the stimulus-evoked oscillatory phenomenon is reliable, its roles in sensation, perception, memory formation and

pattern recognition remain to be demonstrated—a task requiring a behavioural paradigm. Using honeybees, we now demonstrate that odour encoding involves, as it does in locusts, the oscillatory synchronization of assemblies of projection neurons and that this synchronization is also selectively abolished by picrotoxin, an antagonist of the GABA_A (γ-aminobutyric acid) receptor. By using a behavioural learning paradigm, we show that picrotoxin-induced desynchronization impairs the discrimination of molecularly similar odorants, but not that of dissimilar odorants. It appears, therefore, that oscillatory synchronization of neuronal assemblies is functionally relevant, and essential for fine sensory discrimination. This suggests that oscillatory synchronization and the kind of temporal encoding it affords provide an additional dimension by which the brain could segment spatially overlapping stimulus representations.

Investigation of olfactory processing in the locust antennal lobe—a functional and morphological analogue of the vertebrate olfactory bulb—has indicated that both monomolecular and complex odours are represented there combinatorially by dynamical assemblies of projection neurons^{5,9–11}. Each neuron in an odour-coding assembly responds with an odour-specific temporal firing pattern consisting of periods of activity and silence^{5,9}. Any two neurons responding to the same odour are usually co-active only during a fraction of the population response. The spikes of co-activated neurons are generally synchronized^{5,9,10} by the distributed action of GABA-ergic local neurons¹², resulting in large-amplitude, 20–35 Hz local field potential (LFP) oscillations in their target area, the calyx of the mushroom body⁵. Each successive cycle of the odour-evoked oscillatory LFP can therefore be characterized by a co-active subset of projection neurons, and an odour is thus represented by a specific succession of synchronized assemblies^{10,11}. This representation thus comprises three main features—the identity of the odour-activated neurons, the temporal evolution of the ensemble, and oscillatory synchronization—whose importance to the animal for learning and recognition needs to be examined.

We have previously shown that picrotoxin (PCT) applied to the locust antennal lobe selectively blocks the fast inhibitory synapse between local and projection neurons and abolishes their oscillatory synchronization: this manipulation altered neither the response profiles of projection neurons to odours, nor their odour specificity¹². We have now made use of this pharmacological tool to assess whether oscillatory synchronization plays a role in odour learning and discrimination, an experiment that requires a behavioural assay. We therefore used honeybees, which can be trained to extend their mouth parts (proboscis) in response to specific odours after a few associative forward pairings of these odours with a sucrose reinforcement (proboscis-extension (PE) conditioning)^{13–15}. First, we demonstrated that odour representation in the honeybee includes the same three features as those discovered in the locust; second, we tested the importance of oscillatory synchronization for odour learning and discrimination.

Odours, but not air alone, puffed onto an antenna of a honeybee evoked bouts of ~30 Hz LFP oscillations in the calyx of the ipsilateral mushroom body (for example, mint; Fig. 1a). These oscillations lasted for ~0.5–1 s in response to a 1-s long odour puff. Sliding-window autocorrelations of these LFPs revealed the sustained periodic structure of the odour-evoked responses (Fig. 1a). Simultaneous intracellular recordings from antennal lobe neurons showed that, as in locusts⁵, individual antennal lobe neurons responded selectively to certain odours with prominent membrane-potential oscillations (Figs 1b, 3a; $n = 21$ neurons in 16 animals) which are locked to the mushroom body LFP (Fig. 1c, d). Mushroom body LFP oscillations lagged behind those in antennal lobe neurons (phase, $-53^\circ \pm 5^\circ$; mean $\pm$ s.e.m.; $n = 290$ cycles, where 0° is defined as the peak of the LFP; Fig. 1c). This is consistent with our findings in locusts, in which LFP oscillations in the mushroom body result, at least in part, from the coherent input

‡ Present address: GICCS, Research Building Room WP-18, Georgetown University Medical Center, 3970 Reservoir Rd, N.W. Washington, DC 20007-2197, USA.

of synchronized and convergent antennal lobe projection neurons^{5,12}. This was directly confirmed with paired intracellular recordings from antennal lobe neurons, showing synchronized membrane-potential oscillations in response to specific odours (Fig. 1e; $n = 2$ animals).

In locusts, the oscillatory synchronization of projection neurons—and thus the resulting odour-evoked mushroom body LFP oscillation—depends on inhibitory feedback from GABA-ergic local neurons¹². This feedback is selectively abolished either by injection into the antennal lobe or by brain superfusion of 100 μ M picrotoxin¹². We therefore applied picrotoxin to the brains of bees and assessed its effect on odour-evoked projection neuron synchronization by assaying mushroom body LFPs (Fig. 2a). Power spectra calculated over three periods (each 1 s long, before, during and after the odour puff) of the LFP waveform showed a typical peak centred on 30 Hz only for the odour period (Fig. 2b,

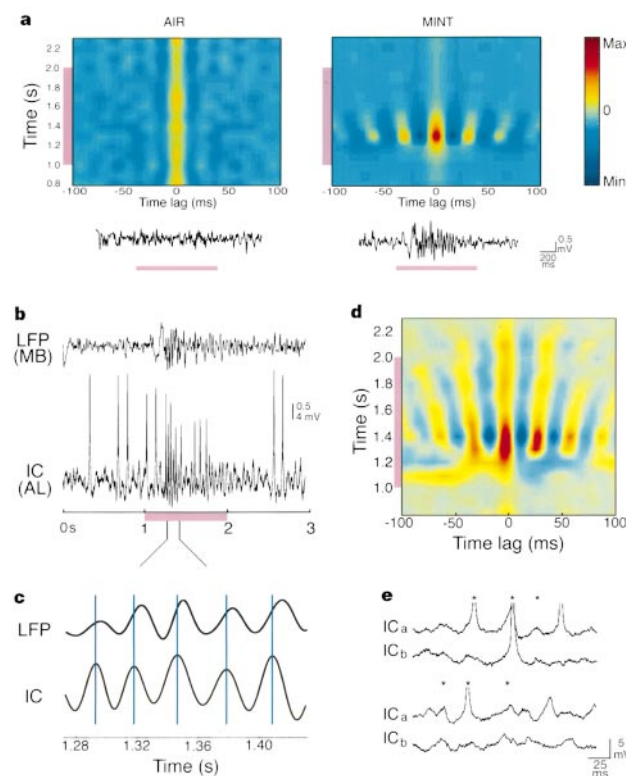


Figure 1 Odours elicit 30-Hz synchronous oscillatory activity in the local field potential (LFP) recorded in the calyx of the mushroom bodies and in antennal lobe neurons. **a**, Mint, but not air, elicits regular oscillations in the LFP (bottom traces), evident also in the repeating banding patterns in sliding-window autocorrelograms⁵ (means of 6 consecutive odour presentations presented at 10-s intervals). Air and mint autocorrelograms are represented on the same scale, indicated by the colour-coded bar. Pink bars indicate odour presentations. **b–e**, Synchronous odour-elicited oscillations were observed in the mushroom body (MB) LFP and in intracellularly (IC) recorded neurons in the ipsilateral antennal lobe (AL). **b**, Mint presentation elicits oscillations in both the neuron and the LFP. The oscillatory LFP indicates rhythmic and synchronized firing of many antennal lobe neurons during the odour response. **c**, Detail of the example in **b** (both records band-pass-filtered at 5–55 Hz; vertical blue lines indicate peaks in intracellular membrane-potential waveform), showing phase locking of the two waveforms with a consistent phase lag. **d**, Sliding-window cross-correlogram between AL neuron membrane potential and mushroom body LFP (average of 6 consecutive odour presentations at 0.1 Hz) reveals consistency, regularity, synchrony and phase locking of oscillations in the two recordings. **e**, Paired intracellular recordings from AL neurons (ICa, ICb) responsive to apple odour, showing transient oscillatory synchronization of their membrane-potential waveforms (asterisks). Spikes are truncated.

left). Eight minutes after application of picrotoxin, however, all odour-evoked power around 30 Hz had been abolished (Fig. 2b, right). This can also be seen from the raw LFP data (Fig. 2a) and from sliding-window autocorrelations of the LFP (Fig. 2c).

The absence of significance power at the stimulus-evoked oscillation frequency might be caused by a silencing, rather than desynchronization, of the antennal lobe projection neurons. We therefore repeated the picrotoxin experiments with intracellular recordings from antennal lobe projection and local neurons (Fig. 3). As in locusts, many antennal lobe neurons responded to odours with stimulus-specific slow temporal patterns superimposed on 30-Hz oscillations (Fig. 3A; $n = 8$); we have not yet examined the information content of individual spike times, as we did for locust projection neurons¹⁰. Picrotoxin changed neither the odour selectivity of these neurons, nor the slow temporal features of their responses ($n = 12$). The neurons shown in Fig. 3B, C, for example, retained their response patterns to odours 10 min after picrotoxin application, which did not significantly alter their average firing rate. We conclude that picrotoxin desynchronizes antennal lobe

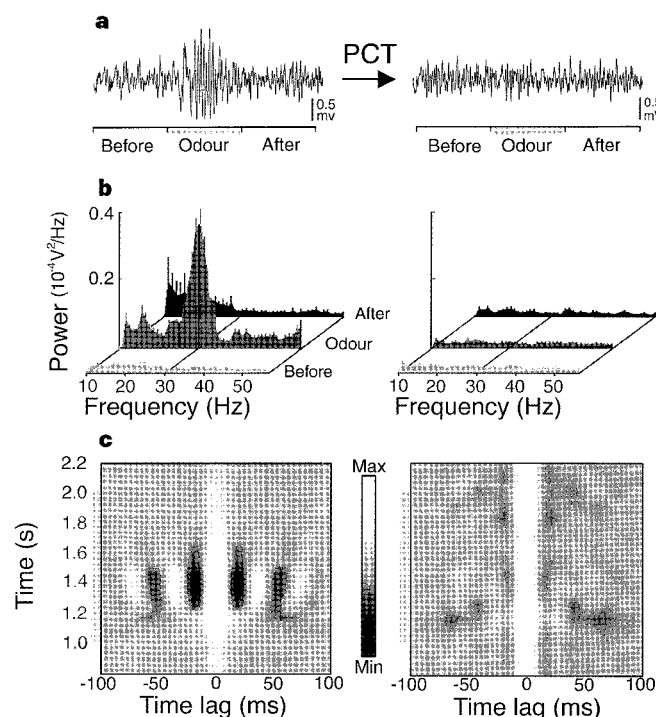


Figure 2 Picrotoxin abolishes the odour-elicited 30-Hz oscillations recorded in the LFP. **a**, Odour (mint) presentation elicits regular, large-amplitude oscillations before (left) but not 8 min after (right) superfusion with PCT. **b**, Power spectra (average of 10 trials at 0.1 Hz; vertical lines indicate s.e.m.) for 1-s periods before, during and after odour presentation, before (left) and 8 min after (right) PCT superfusion. The odorant uniquely and consistently elicits strong 30-Hz LFP oscillations that are abolished by PCT. **c**, Sliding-window autocorrelograms of odour-response periods (mean of 10 consecutive trials at 0.1 Hz) reveal the oscillatory odour response (left) and its suppression after PCT superfusion (right).

projection neurons in response to odours without otherwise altering their individual response patterns, even when these patterns include periods of reduced firing, as observed in locusts¹². Our results establish picrotoxin as a selective pharmacological tool for testing the role of oscillatory synchronization in odour learning and recognition.

We used a proboscis-extension conditioning assay to test whether picrotoxin could disrupt olfactory discrimination^{14,16–18}. When forager honeybees experience forward pairing of an odour (conditioned stimulus) with sucrose reinforcement, their PE response to that odour increases dramatically for 48 h or longer. This increase is due to associative learning mechanisms¹⁷. Typically, the conditioned response generalizes to some extent to odours that are structurally similar to the conditioned odorant¹⁶; for example, once conditioned to an aliphatic alcohol (such as 1-hexanol), bees show a heightened response to structurally similar alcohols (such as 1-octanol). This generalization response is never as strong as the response to the conditioned stimulus itself, but it is higher than the generalization response to structurally dissimilar odorants (such as terpenes).

If oscillatory synchronization plays a role in odour learning or discrimination, application of picrotoxin to the antennal lobe should diminish the ability of animals to discriminate between odours; hence, these animals should show stronger generalization to

odours not experienced during conditioning. Bees were divided into a control group (saline-treated, $n = 36$) and a test group (PCT-treated, $n = 33$) (see Methods). They were individually treated with saline or picrotoxin and then conditioned with C, an aliphatic alcohol, after a recovery interval $t_1 = 10$ min (Fig. 4a). Training and testing were conducted blind. Bees in both PCT- and saline-treated groups learned the conditioned-stimulus sucrose pairing equally well, showing a maximum response by conditioning trial 5 or 6 (Fig. 4b). After a retention interval $t_2 = 90$ min, the two groups were tested with the conditioned stimulus (C), a similar odour (S, an aliphatic alcohol of different chain length) and a dissimilar odour (D, the terpene geraniol). The percentage of animals in each group that responded with a proboscis extension to C, S or D was then measured. Saline-treated bees responded significantly more often to C than they did to S ($P < 0.05$), indicating that they could discriminate between the two related odours (Fig. 4c). PCT-treated bees, by contrast, failed to differentiate C from S (NS; see Fig. 4c). When tested with odour D, animals from both groups performed equally well (that is, each group had an equally low probability of proboscis extension in response to D relative to C; $P < 0.01$; Fig. 4c), indicating that the picrotoxin-injected bees did not have a non-specific learning, memory or performance deficit. Rather, picrotoxin selectively impaired discrimination between C and S. To

Table 1 Percentage of bees in saline- and PCT-injected groups showing PE response

	$t_1 = 45$ min ($n = 40$ per group)		$t_1 = 45$ min ($n = 40$ per group)		$t_1 = 60$ min ($n = 35$ per group)	
	Saline	PCT	Saline	PCT	Saline	PCT
Conditioned stimulus (C)	53	40	29	20	26	26
Similar odour (S)	23*	30 (NS)	9*	14 (NS)	11*	9*
Dissimilar odour (D)	ND	ND	6*	3*	0**	3**

The percentage is shown of animals in saline- and PCT-injected groups that gave a PE response during tests with C, S and D, $t_2 = 60$ min after conditioning. Saline or PCT were administered by antennal lobe injection t_1 min before conditioning (see Methods).

*, ** Significantly different from test with C at $P < 0.05$ or $P < 0.01$, respectively; one-tailed test criteria. ND, not determined; NS, not significant.

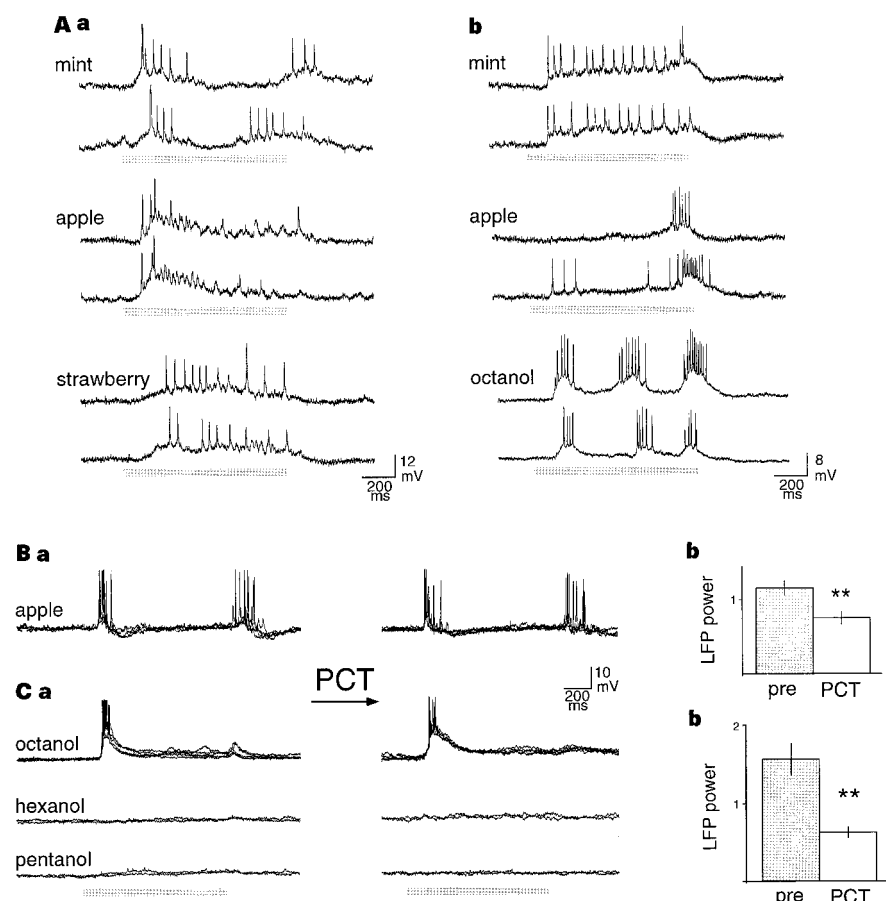


Figure 3 Many AL neurons display odour-specific temporal response patterns. Picrotoxin spares these slow temporal characteristics while abolishing 30-Hz oscillations, as revealed by simultaneous LFP and intracellular recordings.

A, **a** and **b**, Examples of slow temporal response patterns evoked by different odours in two representative antennal lobe neurons (two trials shown for each odour). Subthreshold oscillatory activity is immediately evident in some neurons (**a**) but not in others (**b**). Note the pattern differences across odours and consistency across trials for each odour. **B**, Response pattern of a third neuron to apple odour (**a**, 4 superimposed traces) remains unaffected by PCT application, although LFP power at 32–37 Hz is greatly reduced (**b**, $t = 3.122$, $P < 0.005$). **C**, Response pattern and odour selectivity of a fourth neuron remains unaffected by PCT application (**a**), although LFP power at 31–36 Hz is greatly reduced (**b**; $t = 4.312$, $P < 0.0007$). There are five superimposed traces for octanol and two for hexanol and pentanol, respectively; LFP power is given as in $10^{-4} \text{ V}^2 \text{ Hz}^{-1}$. Some spikes are truncated in **B**, **C**.

determine the time course of picrotoxin activity, we carried out experiments with three other groups ($n = 220$ animals), using different intervals t_1 and t_2 (Table 1). These experiments confirmed the results shown in Fig. 4 and indicated that the behavioural effects of picrotoxin wore off as time elapsed between drug injection and conditioning; for $t_1 = 60$ min and $t_2 = 60$ min, for example, C and S could again be discriminated ($P < 0.05$; Table 1). This recovery time course indicates that the effect of picrotoxin in the experiment shown in Fig. 4c was limited to the conditioning period, the time of specific odour-reinforcement association.

Because picrotoxin affected the behavioural discrimination of similar alcohols, this effect could simply be due to a picrotoxin-induced loss of odour specificity (rather than desynchronization) of the projection neurons representing these alcohols¹⁹. We therefore tested the 'tuning' of antennal lobe neurons to three alcohols (pentanol, hexanol and octanol) before and after picrotoxin application ($n = 9$ animals): in no case did picrotoxin alter the specificity of neuronal tuning, despite significantly reducing odour-induced LFP power around 30 Hz (Fig. 3c).

We have shown that odours evoke the oscillatory synchronization of groups of projection neurons in the honeybee antennal lobe, and that picrotoxin can, as in locusts, abolish oscillatory synchronization while sparing neural response and odour selectivity. Behavioural experiments combining picrotoxin-induced desynchronization of the projection neurons during proboscis-extension conditioning with odour discrimination tests showed that picrotoxin-treated animals failed to discriminate between similar odorants (1-hexanol and 1-octanol) although they could still discriminate between structurally dissimilar ones (that is, either alcohol from the terpene geraniol). These results indicate that neural synchronization—and thus the ability it affords to encode stimulus features in time^{10,11}—plays a role in fine sensory discrimination tasks which require the separation of stimuli whose neural

representations spatially overlap. Synchronization appears to be unimportant, however, for the discrimination of unrelated stimuli (ones less likely to have overlapping neural representations, as suggested by imaging experiments²⁰). Because the effect of PCT was limited to the conditioning period, and because discrimination of dissimilar odours was possible 60–90 min after conditioning, we conclude that PCT did not impair odour learning *per se*. Rather, PCT (and thus desynchronization) appeared to impair the separation of the neural representations of two related odour stimuli, of which one was stored in memory in a form that lacked its natural periodicity and synchronization features. Synchronization of neuronal assemblies therefore appears to be important to help reduce the overlap between the neural representations of related stimuli, possibly by using the temporal aspects of their representations^{10,11} as separable features. It is tempting to speculate that neural oscillatory synchronization might play a similarly important role for refined stimulus encoding and recognition in the other sensory systems and animals where oscillations occur^{1,3,6,8,21,22}. □

Methods

Our results represent experiments conducted with 739 animals (*Apis mellifera*) (450 for physiology, 289 for behaviour).

Preparation for physiology. Foraging worker bees from a colony established by a multiply mated queen were collected as they returned to the hive and were immobilized in a moulded wax holder. The first antennal segment was held forward with a small drop of epoxy resin on its proximal joint. To stabilize the brain for intracellular recording, the oesophagus was gently retracted through a small window cut open between the antennae and the mouth. The oesophagus was held taut and the window was sealed with a drop of wax²³. A second, larger window was then opened posterior to the antennae. Glandular and sheath material were gently removed as the brain was superfused with oxygen-saturated saline (in mM: 140 NaCl, 5 KCl, 5 CaCl₂, 4 NaHCO₃, 1 MgCl₂, 6.3 HEPES, pH 7.0).

Odour stimulation. Controlled puffs of odorant (1 s duration, 0.31 min⁻¹) were delivered by 15 stainless steel nozzles placed 21 mm in front of the antenna. The circularly arrayed nozzles were angled inwards to converge on the antenna. Clean, dry 'background' air was delivered constantly (0.31 min⁻¹). Each odour (3–10 μ l apple, strawberry (Gilberties), cherry (Bell Flavors and Fragrances), spearmint (Flavco), eugenol, geraniol, 1-pentanol, 1-hexanol, 1-octanol (Sigma), cineole, isoamylacetate, citral (Aldrich)) was carried in a separate nozzle; odorants were then vented through an exhaust funnel or a fume hood.

Electrophysiology. Local field potentials were recorded using saline-filled blunt glass micropipettes (tip, $\sim 10 \mu$ m) and were amplified with a d.c. amplifier (NPI, Adams-List). Recordings from antennal lobe neurons were made intracellularly using sharp glass micropipettes (120 M Ω) filled with 0.5 M potassium acetate, and were amplified with a separate d.c. amplifier (Axon). Micropipettes were stretched by a horizontal puller (Sutter). Data were stored on digital tape (Micro Data) and analysed off-line using National Instruments NBMI16L hardware and LabVIEW (National Instruments) and MatLab (The MathWorks) software. Non-phase shifting, band-pass filtering (5–55 Hz, 5-pole; Butterworth) was accomplished by using a software algorithm. Picrotoxin (PCT, Sigma; 100 μ M in oxygen-saturated saline¹²) was superfused over the brain. Cross-correlation analysis and display were carried out as described^{5,9}. Sliding-window correlations were averages of correlations calculated for single trials. Statistical comparisons were made by unpaired two-tailed *t*-tests.

Proboscis extension (PE) conditioning. Bees were individually collected in vials at the entrance of a colony as they returned from or departed on a foraging trip. Vials were immediately placed in an ice-water bath to anaesthetize the animals. Each bee was placed in a harness made of a small metal tube and a strip of tape was inserted dorsally between its head and thorax¹⁵. Details of the PE conditioning protocol are described elsewhere^{24,25}. After a recovery period, bees were tested for motivation by touching one antenna with a droplet of sucrose. Bees that vigorously extended their proboscs to this stimulus were selected for treatment and conditioning. PCT or saline was applied in two ways. In the first method (providing the best learning performance; group 1, Fig. 4), a 3- μ l drop of saline or of 100 μ M PCT in saline was topically applied to the dorsal anterior

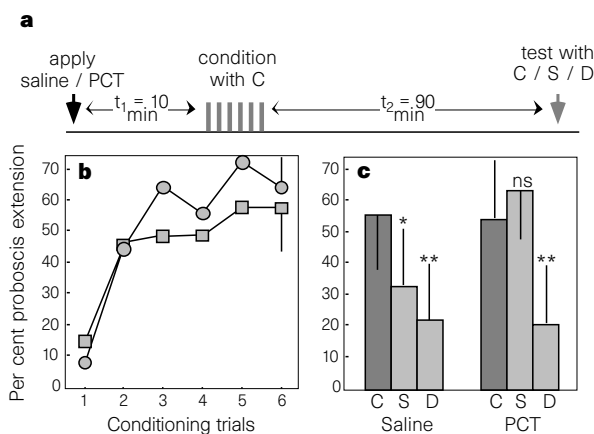


Figure 4 Application of picrotoxin, but not saline, to the antennal lobes impairs discrimination of structurally similar simple odours but not of dissimilar ones. **a**, Experimental protocol. **b**, Acquisition for groups treated with saline (circles; 36 animals) or PCT (squares; 33 animals) was equally rapid and reached asymptote by trial 5. The ordinate represents the percentage of animals that responded with a proboscis extension on each conditioning trial. Vertical lines in **b**, **c** represent 95% confidence intervals. To avoid clutter, we show only the limit (upper or lower) that is relevant for hypothesis testing. **c**, Discrimination of odour C (dark bars) from S differed in groups treated with saline and PCT (same animals as in **b**). Significantly more saline-treated bees responded to C than to S ($\chi^2 = 3.7$, $P < 0.05$; single asterisk). By contrast, PCT-treated bees gave statistically comparable responses to C and S ($\chi^2 = 0.6$; NS). The response levels to C in either group did not differ significantly from one another ($\chi^2 = 0.4$; NS). Discrimination of C from the dissimilar odour (D), however, was possible even in the PCT-treated group (saline group: $\chi^2 = 8.6$, $P < 0.01$, double asterisks; PCT group: $\chi^2 = 8.0$, $P < 0.01$).

surface of the antennal lobes through a small window cut in the head cuticle; these experiments were done without the experimenter knowing whether the drop contained PCT or saline. In the second method (group 2, Table 1), 0.1 nl saline or picrotoxin (100 μ M–1 mM in saline) was injected directly into the antennal lobes through a small window in the head just above the base of each antenna using a Picospritzer (General Valve)²⁶. Injections gave the same results as topical applications, although PE response rates were reduced, as commonly observed after extensive surgery. After a time t_1 (10, 45, 60 or 90 min) for recovery, animals were trained by using the following protocol^{13,14}: 6 paired presentations of odorant (4-s pulse into a vented air stream) and sucrose (0.4 μ l of 1.25 M solution for group 1, 2 μ l of 2 M solution for group 2, presented to the antenna and the proboscis 3 s after odorant pulse onset), every 2 min (group 1) or 30 s (group 2). Animals showing a PE response in each trial were selected to receive 2 or 3 extinction (odour only) trials (one with each of the 2 or 3 test odours; see below) 90 min (group 1) or 60 min (group 2) after conditioning. The odourants used for conditioning were 1-hexanol or 1-octanol. Groups were counterbalanced to contain roughly equal numbers of bees trained with either alcohol. The odours used for testing (1-octanol, 1-hexanol, geraniol) were presented to each animal in a randomized order. Generalization between the alcohols and geraniol is typically low¹⁶. We used the percentage of subjects that responded to an extinction test as the response measure. Results were compared with χ^2 statistics because behavioural data were categorical (PE or no PE). Statistical values are one-tailed because generalization responses were not expected to exceed the response levels to conditioned stimuli.

Received 9 July; accepted 6 August 1997.

- Adrian, E. D. Olfactory reactions in the brain of the hedgehog. *J. Physiol. (Lond.)* **100**, 459–473 (1942).
- Gray, C. M. & Skinner, J. E. Centrifugal regulation of neuronal activity in the olfactory bulb of the waking rabbit as revealed by reversible cryogenic blockade. *Exp. Brain Res.* **69**, 378–386 (1988).
- Gelperin, A. & Tank, D. W. Odour-modulated collective network oscillations of olfactory interneurons in a terrestrial mollusc. *Nature* **345**, 437–440 (1990).
- Delaney, K. R. *et al.* Waves and stimulus-modulated dynamics in an oscillating olfactory network. *Proc. Natl Acad. Sci. USA* **91**, 669–674 (1994).
- Laurent, G. & Davidowitz, H. Encoding of olfactory information with oscillating neural assemblies. *Science* **265**, 1872–1875 (1994).
- Gray, C. M. & Singer, W. Stimulus specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl Acad. Sci. USA* **86**, 1698–1702 (1989).
- Singer, W. & Gray, C. M. Visual feature integration and the temporal correlation hypothesis. *Annu. Rev. Neurosci.* **18**, 555–586 (1995).
- Neuenschwander, S. & Varela, F. J. Visually triggered neuronal oscillations in the pigeon—an autocorrelation study of tectal activity. *Eur. J. Neurosci.* **7**, 870–881 (1993).
- Laurent, G., Wehr, M. & Davidowitz, H. Temporal representations of odors in an olfactory network. *J. Neurosci.* **16**, 3837–3847 (1996).
- Wehr, M. & Laurent, G. Temporal combinatorial encoding of odours with oscillations. *Nature* **384**, 162–166 (1996).
- Laurent, G. Dynamical representation of odors by oscillating and evolving neural assemblies. *Trends Neurosci.* **19**, 489–496 (1996).
- MacLeod, K. & Laurent, G. Distinct mechanisms for synchronization and temporal patterning of odor-encoding neural assemblies. *Science* **274**, 976–979 (1996).
- Kuwabara, M. Bildung des bedingten reflexes von pavlovs typus bei der honigbiene, *Apis mellifica*. *J. Fac. Sci. Hokkaido Univ. (Ser. VI Zool.)* **13**, 458–464 (1957).
- Bitterman, M. E., Menzel, R., Fietz, A. & Schäfer, S. Classical conditioning of proboscis extension in honeybees (*Apis mellifera*). *J. Comp. Psychol.* **97**, 107–119 (1983).
- Menzel, R. & Bitterman, M. E. In *Neuroethology and Behavioral Physiology* (eds Huber, F. & Markl, H.) 206–215 (Springer, New York, 1983).
- Smith, B. H. & Menzel, R. The use of electromyogram recordings to quantify odorant discrimination in the honey bee, *Apis mellifera*. *J. Insect Physiol.* **35**, 369–375 (1989).
- Menzel, R. In *Neurobiology of Comparative Cognition* (eds Kesner, R. P. & Olton, D. S.) 237–292 (Erlbaum, New Jersey, 1990).
- Menzel, R., Michelson, B., Rüffer, P. & Sugawa, M. In *Modulation of Synaptic Plasticity in Nervous Systems* NATO ASI series, Vol. H19 (eds Hertung, G. & Spatz, H.-C.) 335–350 (Springer, Berlin, 1988).
- Yokoi, M., Mori, K. & Nakanishi, S. Refinement of odor molecule tuning by dendrodendritic synaptic inhibition in the olfactory bulb. *Proc. Natl Acad. Sci. USA* **92**, 3371–3375 (1995).
- Joerges, J., Kuttner, A., Galizia, C. G. & Menzel, R. Representation of odours and odour mixtures visualized in the honeybee brain. *Nature* **387**, 285–288 (1997).
- Murthy, V. N. & Fetis, E. E. Coherent 25 Hz to 35 Hz oscillations in the sensorimotor cortex of awake behaving monkeys. *Proc. Natl Acad. Sci. USA* **89**, 5670–5674 (1992).
- Gray, C. M. Synchronous oscillations in neuronal systems: mechanisms and functions. *J. Comput. Neurosci.* **1**, 11–38 (1994).
- Mauelshagen, J. Neural correlates of olfactory learning-paradigms in an identified neuron in the honeybee brain. *J. Neurophysiol.* **69**, 609–625 (1993).
- Bhagavan, S. & Smith, B. H. Olfactory conditioning in the honey-bee, *Apis mellifera*—effects of odor intensity. *Physiol. Behav.* **61**, 107–117 (1997).
- Smith, B. H. An analysis of blocking in binary odorant mixtures: An increase but not a decrease in intensity of reinforcement produces unblocking. *Behav. Neurosci.* **111**, 57–69 (1997).
- Macmillan, C. S. & Mercer, A. R. An investigation of the role of dopamine in the antennal lobes of the honeybee, *Apis mellifera*. *J. Comp. Physiol. A* **160**, 359–366 (1987).

Acknowledgements. We thank K. MacLeod, L. Kay, M. Wehr, A. Hershowitz and H. Krapp for their helpful comments. Supported by an NRSA (NIDCD) fellowship (M.S.), an NIMH grant (B.H.S.), an NSF grant, an NSF Presidential Faculty Fellow award, and a grant from the Sloan Center for Theoretical Neuroscience at Caltech (G.L.).

Correspondence and requests for materials should be addressed to G.L. (e-mail: laurentg@cco.caltech.edu).

Prion (PrP^{Sc})-specific epitope defined by a monoclonal antibody

C. Korth*, B. Stierli†, P. Streiti†, M. Moser*, O. Schaller*, R. Fischer‡, W. Schulz-Schaeffer§, H. Kretschmar§, A. Raeber||, U. Braunf, F. Ehrensperger☆, S. Hornemann#, R. Glockshuber#, R. Riek#, M. Billeter#, K. Wüthrich# & B. Oesch*

* Prionics AG, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

† Brain Research Institute, University of Zürich, 8029 Zürich, Switzerland

‡ Institut für Biochemie, ETH Zürich, 8092 Zürich, Switzerland

§ Institut für Neuropathologie, Universität Göttingen, 37075 Göttingen, Germany

|| Institut für Neuropathologie, University of Zürich, 8091 Zürich, Switzerland

f Klinik für Wiederkäufer- und Pferdemedizin, University of Zürich, 8057 Zürich, Switzerland

☆ Institut für Veterinärpathologie, University of Zürich, 8057 Zürich, Switzerland

Institut für Molekularbiologie und Biophysik, ETH Zürich, 8093 Zürich, Switzerland

Prions are infectious particles causing transmissible spongiform encephalopathies (TSEs). They consist, at least in part, of an isoform (PrP^{Sc}) of the ubiquitous cellular prion protein (PrP^C). Conformational differences between PrP^C and PrP^{Sc} are evident from increased β -sheet content and protease resistance in PrP^{Sc} (refs 1–3). Here we describe a monoclonal antibody, 15B3, that can discriminate between the normal and disease-specific forms of PrP. Such an antibody has been long sought as it should be invaluable for characterizing the infectious particle as well as for diagnosis of TSEs such as bovine spongiform encephalopathy (BSE) or Creutzfeldt–Jakob disease (CJD) in humans. 15B3 specifically precipitates bovine, murine or human PrP^{Sc}, but not PrP^C, suggesting that it recognizes an epitope common to prions from different species. Using immobilized synthetic peptides, we mapped three polypeptide segments in PrP as the 15B3 epitope. In the NMR structure of recombinant mouse PrP, segments 2 and 3 of the 15B3 epitope are near neighbours in space, and segment 1 is located in a different part of the molecule. We discuss models for the PrP^{Sc}-specific epitope that ensure close spatial proximity of all three 15B3 segments, either by intermolecular contacts in oligomeric forms of the prion protein or by intramolecular rearrangement.

PrP-null mice were immunized with full-length recombinant bovine PrP. After fusion of spleen cells with myeloma cells, we selected ~50 hybridoma cells that produced monoclonal antibodies recognizing either native bovine PrP^{Sc} (PrP^{BSE}) immobilized on nitrocellulose or recombinant bovine PrP (rbPrP) in an enzyme-linked immunosorbent assay (ELISA). One of these antibodies (15B3) was selected for binding to protease-digested BSE brain homogenates; a second (6H4) efficiently recognized recombinant PrP. On western blots, 6H4 recognized rbPrP, as well as bovine, human, mouse and sheep PrP^C, whereas 15B3 did not react with any form of PrP (results not shown). To determine the reactivity of these antibodies with native PrP^C and PrP^{Sc}, we immunoprecipitated PrP from brain homogenates of normal and BSE-infected cattle. The precipitated proteins were then analysed on western blots using a rabbit polyclonal antiserum to rbPrP (Fig. 1). The 6H4 antibody precipitated PrP from BSE as well as from normal brain homogenates; 15B3 precipitated only PrP from brain homogenates of BSE-diagnosed cattle (Fig. 1a). Upon proteinase K treatment, normal PrP is completely digested, whereas the 33K–35K form of PrP^{Sc} is shortened to 27K–30K (PrP 27–30), probably as a result of

degradation of the amino-terminal segment of residues 23–90, analogous to hamster PrP^{Sc} (ref. 3). Digestions of brain homogenates or immunoprecipitates with proteinase K are shown in Fig. 1b. Proteinase K digestion of BSE homogenates or the immunoprecipitate with 15B3 yields the PrP^{BSE}-specific band of 27K–30K (Fig. 1b). Not all of the immunoprecipitated PrP was protease-resistant, suggesting that 15B3 recognizes multiple forms of disease-specific PrP with different sensitivities to proteinase K. Apparently, PrP with properties characteristic for PrP^{Sc} but without protease resistance occurs as an intermediate in the generation of fully proteinase-resistant PrP^{Sc} (ref. 4). No PrP 27–30 was found in normal homogenates or immunoprecipitates with protein A only (Fig. 1b) or 6H4 (not shown). 15B3 therefore seems to be a PrP^{BSE}-specific antibody, even though we immunized with recombinant bovine PrP. Injection of rbPrP into Tg20 mice overexpressing mouse PrP^{Sc} has not produced disease for 430 days, whereas two out of four mice injected with a homogenate from the medulla of a BSE-affected cow have come down with TSE at 388 and 426 days (A.R., C.K. and B.O., unpublished results). In addition, it has been reported that recombinant PrP is not infectious⁶. Recombinant PrP is also not protease-resistant, which is a hallmark of PrP^{Sc} (C.K., unpublished observation)^{7,8}. The model of Lansbury and Caughey⁹, which postulates that the two isoforms of PrP are in a dynamic equilibrium, provides a possible explanation for these findings. By immunizing with large amounts of normal PrP, a small portion of the protein might, according to this hypothesis, have been in the scrapie-specific conformation when triggering the immune response. Alternatively, recombinant PrP molecules might transi-

ently associate (see below and Fig. 3b), and thereby form the prion-specific epitope when acting as an immunogen.

We further analysed the species specificity of 15B3 using mouse scrapie-infected brain homogenates (Fig. 1c) and brain homogenates from CJD type-1 patients (Fig. 1d)¹⁰. For comparison, the mouse brain homogenates of PrP-null as well as normal and scrapie-infected wild-type mice, and the immunoprecipitates corresponding to twice the amount of the homogenates are shown (Fig. 1c). Mouse PrP^{Sc} could be efficiently precipitated by 15B3, as indicated by the presence of PrP 27–30 after the digestion with proteinase K (Fig. 1c, lane a). When brain homogenate was treated with proteinase K before the precipitation, 15B3 was also able to precipitate PrP 27–30 (Fig. 1c, lane b), indicating that the N-terminal segment 23–90 is not critical for binding of 15B3 to PrP^{Sc}, even though precipitation of intact PrP^{Sc} appeared to be more efficient than that of PrP 27–30. Proteinase K digestion causes the formation of large aggregates (scrapie-associated fibrils) which may mask the 15B3 epitope. Surprisingly, 15B3 also specifically recognized PrP^{CJD} from sporadic CJD cases but not human PrP^C (Fig. 1d), even though the amino-acid sequence in the regions of the 15B3 epitope is not fully conserved (see below, and Fig. 2b).

The epitopes recognized by the two antibodies were determined by using a gridded array of synthetic peptides consisting of 104 13-residue peptides sequentially shifted in steps of two amino acids and covering the whole mature bovine PrP sequence. A single linear epitope (DYEDRYRE; corresponding to positions 144–152 of human PrP¹¹) was mapped for 6H4, whereas three distinct peptide sequences were found to form the 15B3 epitope (amino acids

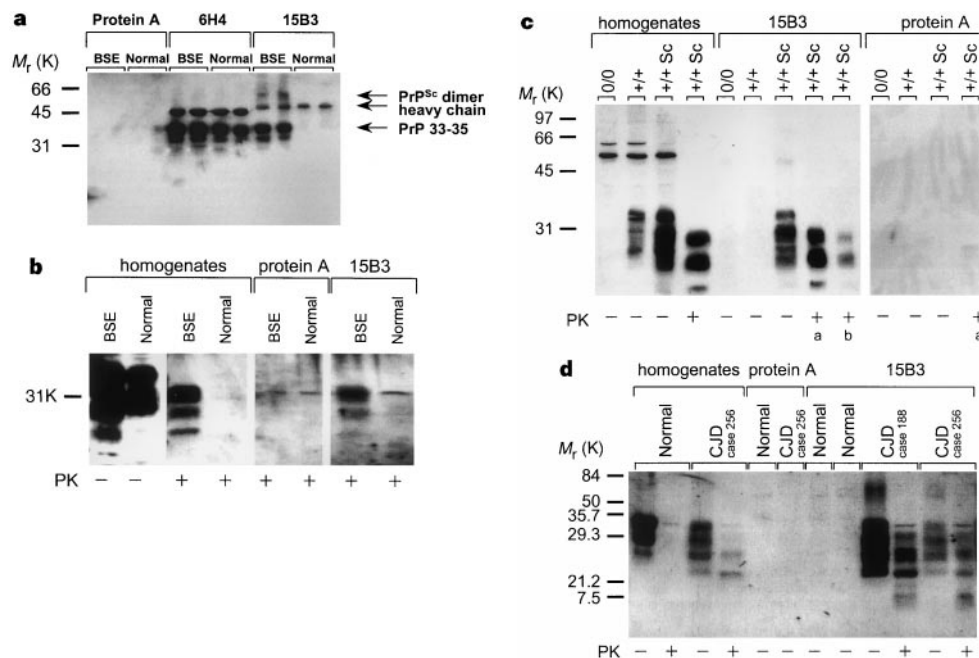


Figure 1 Immunoprecipitation of bovine, mouse and human PrP with monoclonal antibodies 15B3 and 6H4. **a**, The supernatant of a centrifuged homogenate from the medulla of two different BSE-diagnosed or two normal animals was incubated with antibodies 6H4 or 15B3. Antibodies were precipitated with protein A- (15B3) or protein G-agarose (6H4). As a control, protein A only was incubated without antibodies. Precipitates were analysed on a western blot for the presence of PrP using a polyclonal rabbit antiserum to bovine PrP and goat-anti-rabbit Ig coupled to alkaline phosphatase. Signals were developed with chemiluminescence substrates. Crossreaction of the secondary antibody with immunoprecipitated mouse immunoglobulins leads to the prominent band at about 50K. Note the 60K band characteristic for PrP^{Sc} in the 15B3 but not in the 6H4 immunoprecipitations²⁵. **b**, Proteinase K digestion of PrP^{BSE} immunoprecipitated with mAb 15B3. Undigested and digested bovine brain homogenates were compared

to proteinase K digested immunoprecipitates with protein A-agarose only or with 15B3. The sharp band at 31K represents a crossreactivity of the secondary antibody with proteinase K. The same immunoprecipitates and method of analysis were used as in **a**, **c**. Immunoprecipitation of mouse PrP^{Sc} with mAb 15B3. Homogenates from PrP-null mice (0/0) or wild-type mice (normal (+/+) or scrapie-infected (+/+Sc)) were immunoprecipitated with mAb 15B3 or protein A-agarose only and analysed by western blotting as described. Digestion with proteinase K after (a) or before (b) the immunoprecipitation is indicated. Detection of PrP was done as described. **d**, Immunoprecipitation of human PrP^{CJD} with mAb 15B3. Brain homogenates (cerebellum) from normal persons or CJD patients type 1 (ref. 10) were immunoprecipitated and analysed as described for **a**. Two representative examples from a total of 4 normal persons and 4 CJD cases are shown.

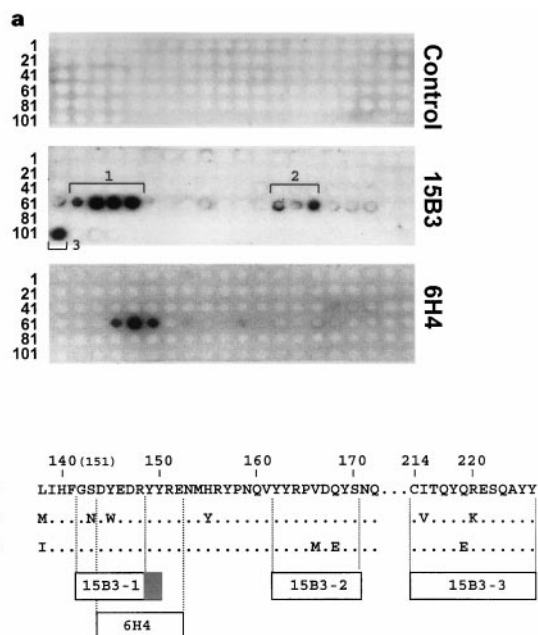


Figure 2 Determination of epitopes for mAbs 15B3 and 6H4. **a**, A gridded array of synthetic peptides corresponding to bovine PrP was incubated with 15B3, 6H4 or with secondary antibody only (peroxidase-labelled goat anti-mouse Ig; control). Bound antibody was visualized with chemiluminescence. Each spot corresponds to a 13-amino-acid peptide, which is shifted by two amino acids along the bovine PrP sequence relative to the previous peptide. Peptides were covalently attached at the C-terminus to a cellulose support. A total of 104 peptides were used to cover the whole bovine PrP sequence including the six octapeptide repeat sequences²¹. **b**, Minimal sequences recognized by 15B3 or 6H4 antibodies in the array of synthetic bovine PrP peptides. The polypeptide segments of the 15B3 epitope are numbered as they occur in the amino acid sequence. The first 15B3 segment extends by two amino acids C-terminally (grey box) if spot number 62, which binds 15B3 only weakly, is excluded. The numbering of the sequence is according to human PrP^{Sc}; the number in brackets indicates the position in the bovine PrP sequences used for the construction of the array of synthetic peptides. Differences with the human and mouse PrP sequences are indicated.

142–148, 162–170, and 214–226; Fig. 2a, b). The relative positions of these partial epitopes in the amino-acid sequence revealed an overlap of the 6H4 epitope with the first segment of the 15B3 epitope (Fig. 2b).

Mapping of the 15B3 epitope onto the NMR structure of the C-terminal domain of mouse PrP (ref. 12) reveals close proximity of the peptide segments 2 and 3, but a much larger spatial separation of the segment 1 from either of the other two components (Fig. 3a). The peptide segment 1 occupies the N-terminal half of helix 1 plus the two residues preceding it, and it is recognized by 6H4 in PrP^C as well as by 15B3 in PrP^{Sc}. This finding would be compatible with either of the two following assumptions: (1) 15B3 recognizes segment 1 of its epitope only in concert with the segments 2 and 3; (2) the polypeptide segment of helix 1 is differently folded in PrP^C and PrP^{Sc}. Component 2 of the 15B3 epitope is in the loop connecting the second β -strand to the second helix, with two thirds of it in a disordered region in the three-dimensional structure, and component 3 is located at the C-terminal end of helix 3 (refs 12, 13). The peptide segments 2 and 3 are located in the proposed binding region for 'protein X' (ref. 14), which is characterized by significant alterations of the electrostatic surface potential among different mammalian species¹⁵: human PrP differs from bovine and mouse PrP in the replacement of the glutamine residues 168 and 219 by glutamic acid residues, as well as by conservative substitutions at positions 166 and 215. As bovine, mouse and human PrP^{Sc} are all precipitated by 15B3 (Fig. 1a–d), this antibody probably binds to the conserved residues in this region.

A single continuous 15B3 binding site could be formed either by aggregation of two or several PrP molecules¹⁶, or by structural rearrangement of a single PrP molecule, or by a combination thereof. Figure 3b suggests a spatial arrangement of a PrP dimer that would bring all three segments of the 15B3 epitope into spatial proximity, with minimal conformational changes of the individual molecules. It is based on the observation of a structural similarity between PrP(121–231) and haemoglobins, which allows a superposition of the helices 1, 2 and 3 of PrP^C onto the helices 1, 6 and 7 of the haemoglobin β -subunit, with a root-mean-square distance for the polypeptide backbone of 2.4 Å. Superposition of two molecules

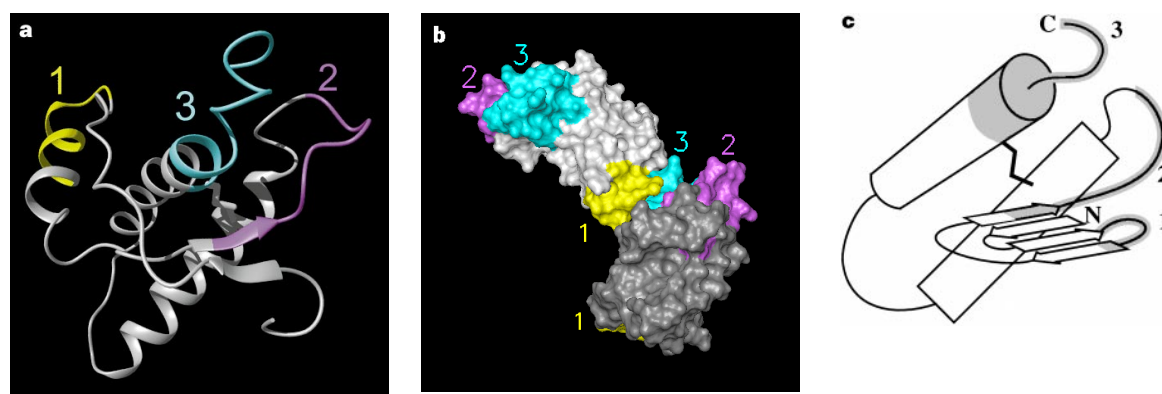


Figure 3 The epitope of the monoclonal antibody 15B3 in the three-dimensional prion protein structure. **a**, Mapping of the 15B3 epitope onto a ribbon drawing of the NMR structure of the C-terminal domain PrP(121–231) of mouse PrP^C (refs 12, 26). The three segments of the 15B3 epitope are coloured yellow (1), violet (2) and cyan (3) in the order in which they occur in the amino acid sequence (Fig. 2b). The residual parts of the molecule and the single disulphide bridge are grey. Regular secondary structures are indicated by ribbons for helices and arrows for β -strands. Drawings in **a** and **b** were prepared with MOLMOL²⁷. **b**, Surface representation of two PrP(121–231) molecules after superposition onto two β -subunits of the crystal lattice of sickle cell haemoglobin¹⁷ (Protein Data Bank entry

1HBS). The segments of the 15B3 epitope are numbered and coloured as in **a**. The superposition included the backbone atoms of residues 145–154, 179–189 and 201–217 of the helices 1, 2 and 3 of PrP(121–231) and of residues 5–14, 106–116 and 125–141 of the helices 1, 6 and 7 of haemoglobin S (r.m.s.d. = 2.4 Å). **c**, Hypothetical fold of the prion protein that would bring all three components of the 15B3 epitope into spatial proximity. The two cylinders represent the disulphide-linked helices 2 and 3 in the orientation of **a**. Helix 1 and the β -sheet have been replaced by four β -strands that form a greek key motif. The chain termini are labelled N and C, and the segments of the 15B3 epitope are numbered as in **a**.

of PrP(121–231) onto two adjacent β -chains in the crystal lattice of haemoglobin S¹⁷ (PDB entry 1HBS) brings the peptide segment 1 of one PrP molecule near to the segments 2 and 3 of the other molecule (Fig. 3b). This superposition aligns residue 6 of the β -chains of sickle cell haemoglobin with Trp at position 145 of the mouse prion protein, which is located in the middle of the epitope segment 1 and is fully exposed to solvent. Mutation of Glu 6 to valine is responsible for the formation of haemoglobin aggregates in sickle cell anaemia.

Intramolecular structural rearrangement bringing all three segments of the 15B3 epitope into close spatial proximity might involve the first helix and lead to an extension of the existing β -sheet¹³. In Fig. 3c, the resulting β -structure is assumed to consist of four strands aligned to form a greek-key motif. Other PrP^{Sc} models that would also lead to close approach of the three segments of 15B3 epitope have been described¹⁸.

The identification of an antibody that binds selectively to PrP^{Sc} from various species provides a new means to identify PrP^{Sc} directly without using proteinase K digestion as a criterion. It will be interesting to see whether 15B3 will be able to neutralize infectivity and thus be a potential therapeutic reagent. The low level of PrP^{Sc} in peripheral tissues has made it difficult to use it as a marker for prion diseases¹⁹. Affinity selection of PrP^{Sc} with 15B3 will allow enrichment of the abnormal isoform of PrP and thus lower the detection limit for PrP^{Sc}, so a prion test for living humans or animals is conceivable. The mapping and three-dimensional modelling of the 15B3 epitopes has provided a view of a prion-disease-specific epitope and may represent a starting point for the production of further diagnostic or therapeutic tools for TSEs. □

Methods

Materials. BSE material was from naturally occurring Swiss cases of BSE, CJD brain material from patients suffering from CJD type 1 (ref. 10), which had been diagnosed using histopathology and immunohistochemistry for PrP. For mouse scrapie material, CD-1 mice were experimentally infected with the RML strain²⁰.

Preparation of recombinant bovine PrP. The bovine PrP open reading frame was amplified by PCR from genomic DNA using the primers 5'-GGGAATTC-CATATGAAGAAGCGACCAAAACCTG and 5'-CGGGATCCTATTAACCTG-CCCCTCGTTGGTA. The resulting PCR product was cloned into pET11a (Novagen) using the *Nde*I and the *Bam*HI restriction sites. The resulting plasmid (pbPrP3) was transfected into *E. coli* BL21(DE3). Bacteria were grown to OD₆₀₀ = 0.8 then induced with 1 mM IPTG and further grown at 30 °C for 3 h. rbPrP corresponding to the mature form of bovine PrP containing six octapeptides²¹ was purified from inclusion bodies after solubilization in 8 M urea, 10 mM MOPS/NaOH, pH 7.0 (UM-buffer), on a CM sepharose column (Pharmacia; UM, 0–0.5 M NaCl gradient) and reverse-phase HPLC (Vydac C₄ column, 0.1% trifluoroacetic acid, 0–60% acetonitrile gradient; C.K. and B.O., unpublished results). Resulting fractions contained either oxidized (elution time, 29 min) or reduced PrP (elution time, 34 min). Usually, CM-Sephadex fractions were oxidized with 1 μ M CuSO₄ for 1 h before purification by reverse-phase HPLC. Purified rbPrP was analysed by mass spectrometry, indicating a protein of the expected mass in which the N-terminal methionine was uncleaved.

Immunization of PrP-null mice. 100 μ g recombinant bovine PrP in Freund's complete adjuvant was injected into PrP null mice (mixed background 129/Sv and C57BL/6)²² subcutaneously, and 21 and 42 d later with the same amount of antigen in Freund's incomplete adjuvant. Mice were boosted intraperitoneally (day 48) and intravenously (day 49) with recombinant PrP dissolved in PBS. At day 50, mice were decapitated and splenocytes fused to myeloma cells as described²³. Supernatants of the resulting hybridoma cell lines were screened both by ELISA with recombinant bovine PrP as antigen and by ELISA²⁴ using native, protease-digested brain homogenate of BSE-diseased cattle. Positive hybridoma cells were subcloned three times.

Characterization of antibodies. Supernatants of selected hybridomas were used to probe bovine PrP in brain homogenates or recombinant PrP on western blots. To determine the epitopes, antibodies were incubated with a gridded array of peptides comprising 104 polypeptides of 13 amino acids, shifted by two

amino acids and covering the entire mature bovine PrP sequence containing six octapeptide repeats²¹. The peptides were covalently attached at their C termini to a cellulose support as individual spots (Jerini Biotech, Berlin). Bound antibody was detected with goat anti-mouse immunoglobulin coupled to horseradish peroxidase and chemiluminescence. Signals were recorded on Hyperfilm ECL (Amersham). For immunoprecipitation, 200 μ l 1% brain homogenates (precleared by centrifugation at 13,000g for 15 min) were incubated for 2 h at room temperature with 200 μ l 0.25 mg ml⁻¹ antibody-containing serum-free medium; after incubation with an additional 50 μ l protein A- or protein G-coupled agarose (for 15B3 and 6H4, respectively; Boehringer Mannheim) for 2 h at room temperature, agarose beads were centrifuged at 13,000g for 3 min and the pellet washed according to the manufacturer. Proteinase K (PK) digestions of immunoprecipitates were done with 20 μ g ml⁻¹ PK (Sigma) for 30 min at 37 °C. Pellets were then boiled in SDS-sample buffer for analysis on western blots. Immunoprecipitated PrP was detected with polyclonal antibodies raised against bovine recombinant PrP in rabbits (R 26) followed by incubation with a goat anti-rabbit immunoglobulin coupled to peroxidase or a goat anti-rabbit IgG coupled to alkaline phosphatase. Bound enzymatic activity was visualized with chemiluminescent substrates (ECL, Amersham, or CSPD, Tropix, respectively).

Received 23 July; accepted 2 October 1997.

- Pan, K. M. *et al.* Conversion of alpha-helices into beta-sheets features in the formation of the scrapie prion proteins. *Proc. Natl Acad. Sci. USA* **90**, 10962–10966 (1993).
- McKinley, M. P., Bolton, D. C. & Prusiner, S. B. A protease-resistant protein is a structural component of the scrapie prion. *Cell* **35**, 57–62 (1983).
- Oesch, B. *et al.* A cellular gene encodes scrapie PrP 27–30 protein. *Cell* **40**, 735–746 (1985).
- Daude, N., Lehmann, S. & Harris, D. A. Identification of intermediate steps in the conversion of a mutant prion protein to a scrapie-like form in cultured cells. *J. Biol. Chem.* **272**, 11604–11612 (1997).
- Fischer, M. *et al.* Prion protein (PrP) with amino-proximal deletions restoring susceptibility of PrP knockout mice to scrapie. *EMBO J.* **15**, 1255–1264 (1996).
- Mehlhorn, I. *et al.* High-level expression and characterization of a purified 142-residue polypeptide of the prion protein. *Biochemistry* **35**, 5528–5537 (1996).
- Weiss, S., Rieger, R., Edenhofer, E., Fisch, E. & Winnacker, E. L. Recombinant prion protein rPrP27-30 from Syrian golden hamster reveals proteinase K sensitivity. *Biochem. Biophys. Res. Commun.* **219**, 173–179 (1996).
- Kaneko, K. *et al.* Molecular properties of complexes formed between the prion protein and synthetic peptides. *J. Mol. Biol.* **270**, 574–586 (1997).
- Lansbury, P. T. & Caughey, B. The chemistry of scrapie infection: implications of the 'ice 9' metaphor. *Chem. Biol.* **2**, 1–5 (1995).
- Parchi, P. *et al.* Molecular basis of phenotypic variability in sporadic Creutzfeldt–Jakob disease. *Ann. Neurol.* **39**, 767–778 (1996).
- Schätzl, H. M., Da Costa, M., Taylor, L., Cohen, F. E. & Prusiner, S. B. Prion protein gene variation among primates. *J. Mol. Biol.* **245**, 362–374 (1995).
- Riek, R. *et al.* NMR structure of the mouse prion protein domain PrP(121–231). *Nature* **382**, 180–182 (1996).
- Glockshuber, R. *et al.* Three-dimensional NMR structure of a self-folding domain of the prion protein PrP (121–231). *Trends Biochem. Sci.* **22**, 241–242 (1997).
- Telling, G. C. *et al.* Prion propagation in mice expressing human and chimeric PrP transgenes implicates the interaction of cellular PrP with another protein. *Cell* **83**, 79–90 (1995).
- Billeter, M. *et al.* Prion protein NMR structure and species barrier for prion diseases. *Proc. Natl Acad. Sci. USA* **94**, 7281–7285 (1997).
- Warwick, J. & Gane, P. J. A model for prion protein dimerization based on alpha-helical packing. *Biochem. Biophys. Res. Commun.* **226**, 777–782 (1996).
- Padian, E. A. & Love, W. E. Refined crystal structure of deoxyhaemoglobin S. II. Molecular interactions in the crystal. *J. Biol. Chem.* **260**, 8280–8291 (1985).
- Huang, Z., Prusiner, S. B. & Cohen, F. E. Scrapie prions: a three-dimensional model of an infectious fragment. *Fold. Design* **1**, 13–19 (1996).
- Kitamoto, T., Mohri, S. & Tateishi, J. Organ distribution of proteinase-resistant prion protein in humans and mice with Creutzfeldt–Jakob disease. *J. Gen. Virol.* **70**, 3371–3379 (1989).
- Chandler, R. L. Encephalopathy in mice produced by inoculation with scrapie brain material. *Lancet* **i**, 1378–1379 (1961).
- Goldmann, W., Hunter, N., Martin, T., Dawson, M. & Hope, J. Different forms of the bovine PrP gene have five or six copies of a short, G-C-rich element within the protein-coding exon. *J. Gen. Virol.* **72**, 201–204 (1991).
- Bueler, H. *et al.* Normal development and behaviour of mice lacking the neuronal cell-surface PrP protein. *Nature* **356**, 577–582 (1992).
- Kennett, R. H. *Monoclonal Antibodies. Hybridomas: A New Dimension in Biological Analysis* (eds Kennett, R. H., McKearn, T. J. & Bechtol, K. B.) 365–367 (Plenum, New York, 1980).
- Oesch, B., Jensen, M., Nilsson, P. & Fogh, J. Properties of the scrapie prion protein: quantitative analysis of protease resistance. *Biochemistry* **33**, 5926–5931 (1994).
- Priola, S. A., Caughey, B., Wehrly, K. & Chesebro, B. A 60-kDa prion protein (PrP) with properties of both the normal and scrapie-associated forms of PrP. *J. Biol. Chem.* **270**, 3299–3305 (1995).
- Riek, R., Hornemann, S., Wider, G., Glockshuber, R. & Wüthrich, K. NMR characterization of the full-length recombinant murine prion protein, mPrP(23–231). *FEBS Lett.* **413**, 282–288 (1997).
- Koradi, R., Billeter, M. & Wüthrich, K. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51–55 (1996).

Acknowledgements. We thank C. Weissmann for discussion and for PrP null mice, and M. Schwab and his group (Brain Research Institute) for support and encouragement at an early stage of this project. This work was supported by grants from the Schweizerische Nationalfonds to B.O. (SPP Biotechnologie), K.W. and R.G., from the Herman Herzer-Foundation, Basel, to B.O., and a fellowship from the Ciba Foundation to M.M.

Correspondence and requests for materials should be addressed to C.K. (e-mail: ckorth@hifo.unizh.ch) or B.O. (e-mail: oesch@hifo.unizh.ch).

The prostaglandin receptor EP₄ triggers remodelling of the cardiovascular system at birth

MyTrang Nguyen*, Todd Camenisch*, John N. Snouwaert*, Elizabeth Hicks*, Thomas M. Coffman†, Page A. W. Anderson‡, Nadia N. Malouf§ & Beverly H. Koller*

* Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7248, USA

† Department of Medicine, Duke University and Durham Veterans Affairs Medical Centers, Durham, North Carolina 27710, USA

‡ Division of Pediatric Cardiology, Duke University Medical Center, Durham, North Carolina 27710, USA

§ Department of Pathology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7525, USA

Survival of newborn placental mammals depends on closure of the ductus arteriosus (DA), an arterial connection in the fetus which directs blood away from the pulmonary circulation and towards the placenta where oxygenation occurs¹. Here we show that morphological changes resulting in closure of the DA in mice are virtually identical to those observed in larger mammals, including humans², and that maintenance of the DA in the open, or patent, state in fetal mice is dependent on prostaglandin synthesis. This requirement is absent in mice lacking the prostaglandin E₂ EP₄ receptor (EP₄(-/-) mice). In EP₄(-/-) mice of the 129 strain, remodelling of the DA fails to occur after birth, resulting in a left-to-right shunt of blood and subsequently in death. This suggests that the neonatal drop in prostaglandin E₂ (refs 3–7) that triggers ductal closure is sensed through the EP₄ receptor. In contrast, 5% of EP₄(-/-) mice of mixed genetic background survive, and selective breeding of these mice leads

to a 21% survival rate, suggesting that alleles at other loci can provide an alternative mechanism for ductal closure.

In about 6 in 10,000 full-term infants, and even more frequently in the premature neonate, the DA fails to close⁸. This allows transmission of high aortic pressures to the pulmonary circuit, which in turn may lead to pulmonary venous hypertension, pulmonary oedema, and congestive heart failure⁹. Permanent closure of the DA depends on a series of structural changes that are independent of the reversible muscular contraction of the DA. Endothelial cells detach from the internal elastic lamina, and the subendothelial region begins to accumulate extracellular matrix material and fluid, initiating the formation of intimal cushions. The internal elastic lamina becomes fragmented, allowing smooth muscle cells to migrate into the subendothelial region, widening the intimal cushions and obstructing the vessel lumen. Eventually the DA degenerates into the fibrous cord referred to as the ligamentum arteriosus^{10,11}.

Examination of the DA in normal mice indicates that, as in other species, the DA is a large muscular artery characterized by numerous layers of immature smooth muscle cells separated by layers of elastin (Fig. 1a). The DA constricts within 30 min of birth, obstructing blood flow, and events leading to permanent closure of the mouse DA are initiated within three hours of birth. As noted in other species, DA closure proceeds from the pulmonary to the aortic end¹⁰ of the vessel (Fig. 1a).

Prostaglandin E₂ (PGE₂) is effective in inhibiting the natural tone

Table 1 Survival of EP₄-deficient mice

Strain*	Parental genotype	Genotype of offspring		
		+/+	+/-	-/-
129	+/- × +/-	68	113	0
MB	+/- × +/-	90	150	4
SMB ₁	-/- × +/-	0	79	2
SMB ₄	-/- × +/-	0	66	14

* 129 mice are a mixture of 129/SvEv and 129/OlaC substrains. MB, mixed background mice, are derived from 129/OlaC, C57BL/6 and DBA/2 mice. SMB, selected mixed background, are derived from EP₄(-/-) MB mice (see Methods).

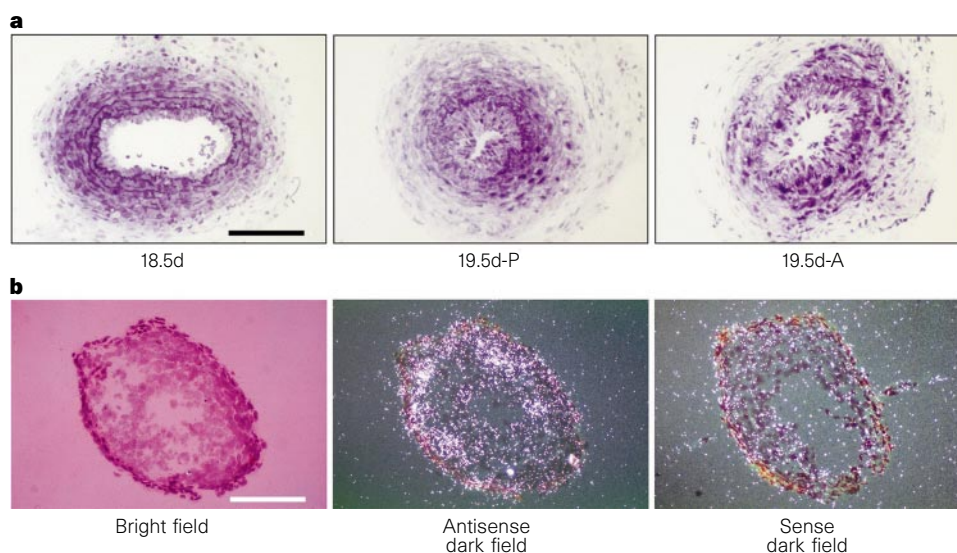


Figure 1 Mouse DA morphology and expression of the EP₄ receptor. **a**, Transverse sections of the ductus arteriosus from fetus at term (18.5d) and neonate, stained with toluidine blue. Sections are taken from the neonatal DA at the region adjacent to the pulmonary trunk (19.5d-P) and towards its junction with the aorta (19.5d-A). The wall of the fetal ductus consists of multiple layers of immature smooth-muscle cells alternating with layers of elastic lamina. This structure alters quickly after birth, as the elastin layer becomes fragmented and the lumen of the vessel is decreased. Projection of the endothelial cells into the lumen, marked oedema, and infiltration of cells into the subendothelial region are

seen. Obstruction of the lumen is less marked in sections of the ductus adjacent to the aorta. Scale bar, 150 μm. **b**, Transverse sections through the DA obtained from a neonate within 12 h of birth were hybridized with ³⁵S-labelled EP₄ antisense probe. Expression of the EP₄ gene is easily detected in the migrating smooth-muscle cells. Bright-field microscopy of the section stained with haematoxylin and eosin is shown on the left. A serial section of the ductus was hybridized with ³⁵S-labelled EP₄ sense probe to establish levels of nonspecific hybridization. Scale bar, 150 μm.

of the DA¹². This property of PGE₂ has been exploited in treating infants with congenital heart defects, because a patent DA can aid in maintaining pulmonary or systemic circulation^{13–15}. To determine whether the mouse DA is also sensitive to PGE₁, PGE₁ was administered to neonatal mice within one hour of birth. In all cases, the DA remained patent in mice that received PGE₁.

Exposure *in utero* to inhibitors of prostanoid synthesis, such as indomethacin, can result in the premature closure of the DA^{16–18}, and a patent DA (PDA) in the premature infant can often be treated with indomethacin^{19,20}. To determine whether prostanoids play a similar role in mice, four pregnant mice were treated with indomethacin and two with vehicle on day 18.5 *post coitus*. Premature closure of the DA was observed in all fetuses from indomethacin-treated dams.

On the basis of pharmacological results, four PGE₂ receptors, designated EP₁, EP₂, EP₃ and EP₄, have been identified, and the genes encoding these subtypes have been cloned²¹. As pharmacological evidence indicates that prostaglandins mediate relaxation of rabbit DA through the EP₄ receptor²², we examined the expression of the EP₄ receptor in mouse DA by *in situ* hybridization. EP₄ expression was easily detected in the smooth muscle cells obstructing the lumen of neonatal DA (Fig. 1b). Expression of EP₄ was also observed in the DA of the preterm 18.5 day fetus (see Supplementary information).

Examination of the DA in mice lacking the EP₄ receptor provides

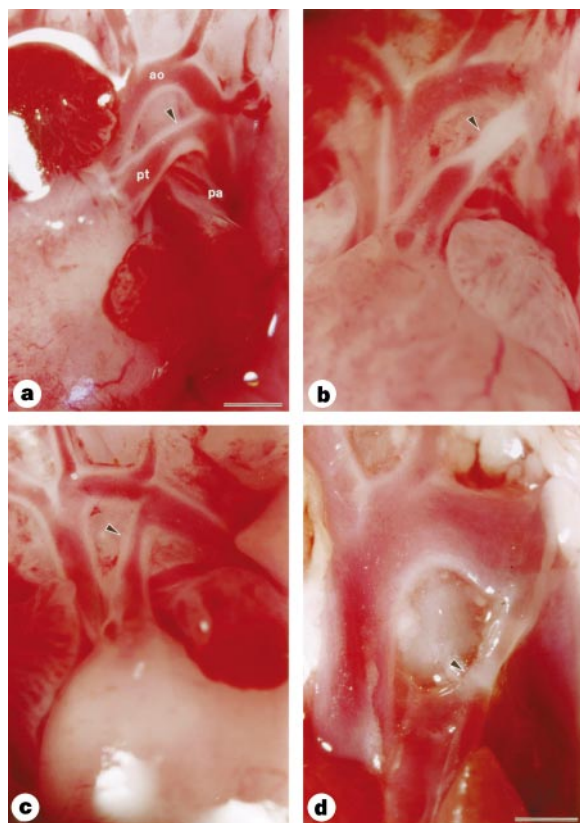


Figure 2 The ductus arteriosus. **a**, From EP₄(+/+) full-term fetus; **b**, from EP₄(+/+) newborn 12 h after birth; **c**, from EP₄(–/–) newborn 12 h after birth; and **d**, from MB EP₄(–/–) animal at 9 weeks old. The pulmonary trunk in the fetus is contiguous with the DA (arrowhead), and connects to the descending aorta at the level of the subclavian artery. In the fetus, the small diameter of the pulmonary artery, a dorsal branch of the pulmonary trunk, reflects minimal blood flow to the lungs before birth. In the EP₄(+/+) animal 2 h after birth (**b**), a lumen is no longer evident in the ductus arteriosus. The pulmonary artery has dilated in comparison to its fetal counterpart. In EP₄(–/–) mice at 12 h (**c**) or 9 weeks (**d**) after birth, the DA remains open. In the 9-week-old animal, however, a zone of constriction (arrow), appears to limit the flow of blood through the shunt. ao, Aorta; pt, pulmonary trunk; pa, pulmonary artery; arrow head, ductus arteriosus. Scale bar, 1 mm.

a direct means for determining the mechanism by which PGE₂ affects the tone and closure of this vessel. A mouse line of the 129 strain carrying a mutant EP₄ was generated by gene targeting in mouse embryonic stem (ES) cells (see Supplementary information). No surviving EP₄(–/–) animals were identified among 181 surviving offspring from EP₄(+/–) crosses analysed at weaning (Table 1). Expected numbers of EP₄(–/–) mice were present in litters genotyped twelve hours after birth but were almost completely absent in 48-hour-old litters.

Although visual examination revealed no obvious differences between newborn EP₄(+/+) and EP₄(–/–) mice, gross examination of internal organs immediately after death revealed that the DA had failed to close in EP₄(–/–) mice (Fig. 2b, c) and was similar in diameter to that of a full-term fetus (Fig. 2a). EP₄(–/–) pups that died had oedema of the lungs. A significantly higher lung-to-body-weight ratio ($P = 1.1 \times 10^{-6}$) was seen on comparison of 10 EP₄(–/–) mice (0.026 ± 0.0009) with 6 controls (0.016 ± 0.0006). Histological analysis of lungs from EP₄(–/–) mice indicated congestion of pulmonary septal capillaries and haemorrhaging into the alveolar spaces and the main respiratory ducts, as well as dilation of perivascular lymphatics (see Supplementary information). We interpret these changes to be the result of a left-to-right shunt caused by the combined effects of a patent DA and postnatal drop in the resistance of the pulmonary circulation.

In some of the EP₄(–/–) pups, particularly those that died more than 24 h after birth, the livers showed changes consistent with passive congestion. Histological analysis also revealed macro- and microvesicular lipid deposits. These changes in the liver are suggestive of acute ischaemia secondary to heart failure.

To verify that pathological changes observed in the lungs and livers of EP₄(–/–) animals were secondary to a patent DA and to determine whether any postnatal changes in the EP₄(–/–) DA could be observed, EP₄(–/–) pups and control littermates were examined at different times. The diameter of the DA was identical in EP₄(–/–) and EP₄(+/+) 18.5-day fetuses. Three hours after birth, the DA was completely closed in EP₄(+/+) mice but not in EP₄(–/–) mice, although the DA in the EP₄(–/–) mice was constricted relative to the fetal DA. Upon examination under magnification, blood could be detected throughout the length of the DA in EP₄(–/–) mice, and patency of the DA in the EP₄(–/–) neonates was verified by applying pressure to the aorta or pulmonary artery and observing blood flow through the DA. No other differences between EP₄(–/–) and EP₄(+/+) mice were detected, either by simple observation or histological examination of tissues at this time. At both 7 and 12 h after birth, the patency of the DA in EP₄(–/–) pups became more obvious. Histological analysis confirmed that the DA of EP₄(–/–) mice remained patent (Fig. 3b, d), whereas in EP₄(+/+) animals (Fig. 3a, c), the vessel had become completely obstructed.

The constriction of the DA in EP₄(–/–) within 3 h of birth suggests that the initial reduction in blood flow through the DA is brought about by prostaglandin-independent contraction of the smooth muscle wall of the vessel. A mechanism in which the natural rise in blood oxygen at birth results in a cytochrome P450-dependent increase in endothelin-1 synthesis has been proposed to mediate these early events in DA constriction²³.

In both dog and humans, the pattern of inheritance of patent DA suggests that it is a multigenetic trait⁸. To determine whether the effect of loss of EP₄ receptors on the closure of the DA can be modified by genes at other loci, chimaeras generated with EP₄(+/–) cells were bred to B6D2 mice (C57BL/6 and DBA/2 mice F₁) to obtain EP₄(+/–) mice on a mixed genetic background. We designated this population of mice mixed-background (MB) strain 129/Ola, C57BL/6, DBA/2. Among the first 244 offspring obtained from the intercross of these EP₄(+/–) mice, four surviving MB-EP₄(–/–) animals were identified (Table 1), suggesting that some combination of genes can support survival of EP₄(–/–) mice. Surviving EP₄(–/–) animals were bred to produce a selected

mixed-background (SMB) population. Five per cent of $EP_4(-/-)$ offspring survive when $EP_4(+/-)$ mice of the MB generation are crossed, and this number is increased to 21% in the SMB_4 offspring.

The DAs of $EP_4(-/-)$ mice that survived were examined, and in all cases closure of the DA was observed. In some MB $EP_4(-/-)$ animals, however, closure was restricted to the region of the DA immediately adjacent to the pulmonary trunk. This produced a funnel-shaped diverticulum that was open at the aortic end (Fig. 2d), similar to the DA diverticulum present in dog strains with susceptibility to patent DA¹⁰. The lack of difference in overall health between surviving $EP_4(-/-)$ mice and control littermates supports the contention that the lung oedema and liver disease seen in the $EP_4(-/-)$ mice that died perinatally was secondary to a patent DA.

We examined the effect of inhibition of prostaglandin synthesis on $EP_4(-/-)$ fetuses by administering indomethacin to $EP_4(+/-)$ females impregnated 18.5 days earlier by MB $EP_4(-/-)$ males. Four hours after treatment, the DA of each pup was examined and its DNA prepared to determine the genotype. In all cases, indomethacin induced closure of the DA in the wild type but not in $EP_4(-/-)$ fetuses (see Supplementary information).

Previous studies indicate that the smooth muscle of the DA has an intrinsic tone that is inhibited by PGE_2 (refs 24, 25). To reconcile these data with those from the genetic experiments presented here, we propose a model in which the EP_4 receptor in the DA acts as a sensor that responds to the perinatal drop in circulating levels of PGE_2 (refs 3–7) by triggering closure of the DA (Fig. 4). In most cells, the EP_4 receptor is coupled to adenylate cyclase through a stimulatory G protein, so that stimulation of the receptor by PGE_2 increases intracellular cyclic AMP levels^{21,26}. We suggest that, following EP_4 receptor expression *in utero*, other pathways contributing to cAMP generation in the cells of the DA are downmodulated so that the EP_4 receptor becomes the principal regulator of intracellular levels of cAMP. In EP_4 -deficient mice, however, patency of the DA fails to become dependent on circulating levels of PGE_2 , explaining why the DA in these animals remains patent after birth and is unresponsive to administration of indomethacin *in utero*.

Genes required for normal closure of the mouse DA, such as the EP_4 receptor, represent candidates for the inherited components of human PDA. An understanding of these genes and the mechanism by which they influence closure of the DA is important, not only for the study and treatment of patent DA, but also for understanding vessel intimal thickening in general and identifying genes that can influence the structure of blood-vessel walls. The DA also provides a model system for unravelling the complex mechanisms by which

PGE_2 and its receptors modulate cell structure and function in response to the changing physiological state of the organism. □

Methods

Tissue preparation for light microscopy. To minimize contraction of the DA in response to exposure to light and oxygen, mouse carcasses were cooled on ice before exposure of the DA. Ducti shown in Fig. 1a were fixed in 2% paraformaldehyde/2.5% glutaraldehyde, embedded in a mixture of Spurr's and Poly-bed 812 resin, and transverse sections, 1 μ m in thickness, were cut and stained with 1% toluidine blue in 1% sodium borate. All other tissue was fixed in 10% phosphate-buffered formalin, pH 7.2, embedded in paraffin, and 5- μ m-thick sections prepared.

RNA *in situ* hybridization. The EP_4 -specific cDNA probe was synthesized by reverse transcription and PCR amplification using oligonucleotide primers EP_4 -1F (5'-GTTTGGCTGATATACTGGTAAAT-3') and EP_4 -2R (5'-ACCTGGTGCTTCATCGACTGGACC-3') and total thymus RNA as a template. These primers correspond to the published EP_4 sequence²⁶. Originally designated as the EP_2 receptor, it is now recognized that this sequence represents the murine homologue of the human EP_4 receptor²⁷. The PCR fragment was cloned using a TA cloning kit (In Vitrogen) and ³⁵S-labelled probes prepared using a commercially available kit (Ambion). The DA was embedded in OCT compound, transverse sections 8 to 10 μ m in thickness were prepared, and *in situ* hybridization carried out as described²⁸.

Generation of EP_4 -deficient mice. The EP_4 cDNA probe was used to isolate genomic clones from a 129/Sv mouse genomic library. Sequence analysis confirmed that these clones corresponded to the mouse EP_4 gene. In the targeting vector, exons 1 and 2 are replaced by the neomycin-resistance gene from the plasmid pJNS2. These exons include the initiation codon and regions of the gene encoding six of the seven transmembrane-spanning domains and putative extracellular loops. 129/Ola derived E14Tg2a ES cells²⁹ were grown, transformed and screened using standard method³⁰. A probe prepared from DNA immediately downstream of the targeted locus was used to identify targeted ES cells and later to genotype the mice generated from these ES cells. Chimaeras derived from targeted ES cells were mated with 129/SvEv or B6D2 (C57BL/6 $\times$ DBA/2 F₁) mice. Mice obtained from crosses with 129/SvEv are here referred to as 129 mice. Analysis of RNA prepared from the thymus and spleen of $EP_4(-/-)$ and $EP_4(+/-)$ newborn pups with EP_4 -specific cDNA probes verified the loss of EP_4 expression in $EP_4(-/-)$ mice. The first $EP_4(-/-)$ mice obtained from crossing $EP_4(+/-)$ offspring of chimaeras and B6D2 animals are designated 'mixed background' (MB). $EP_4(-/-)$ obtained from the crossing of MB $EP_4(-/-)$ mice with either $EP_4(+/-)$ sibs or offspring are designated selected mixed background generation one (SMB_1). $EP_4(-/-)$ mice derived from crosses of $SMB_1EP_4(-/-)$ mice and $EP_4(+/-)$ sibs and offspring are designated SMB_2 .

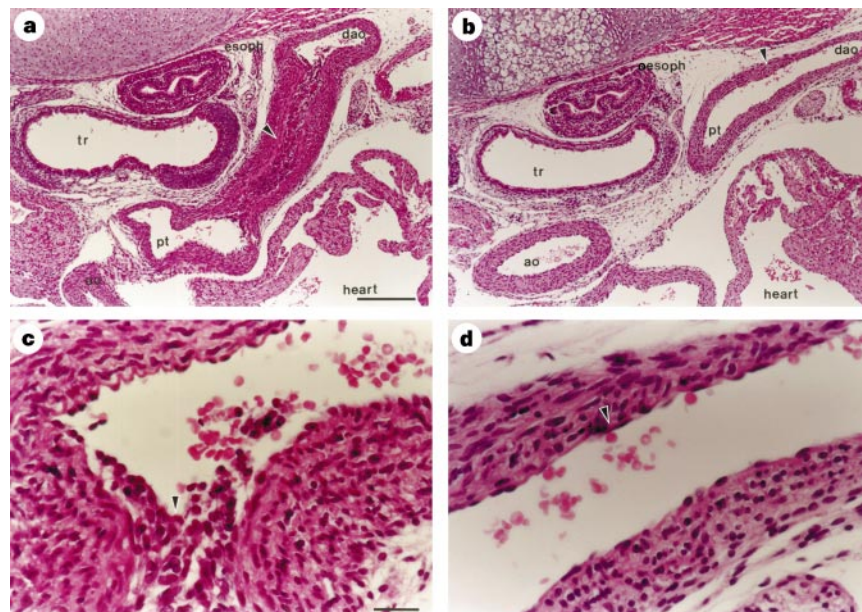


Figure 3 Histological examination of the DA in wild-type and $EP_4(-/-)$ mice. **a**, Twelve hours after birth, the walls of the DA (arrow) of the normal mouse have thickened as a result of the swelling of the intima, and a loose network of cells can be seen filling and obliterating the lumen. **c**, Higher magnification of the junction between the aorta and the ductus shows that these changes end abruptly within the aortic lumen. **b**, In an EP_4 -deficient animal, the ductus remains patent (arrowhead). **d**, Higher magnification of this same section shows the thick wall of the ductus with a clear lumen. Swelling of the intima is not present, and the endothelium remains a single layer of cells contiguous with that of the aorta. oesoph, Oesophagus; tr, trachea; dao, descending aorta; ao, aorta; pt, pulmonary trunk. Arrowhead indicates the ductus arteriosus. Sections were stained with haematoxylin and eosin. Scale bars, 100 μ m (**a** and **b**) or 25 μ m (**c** and **d**).

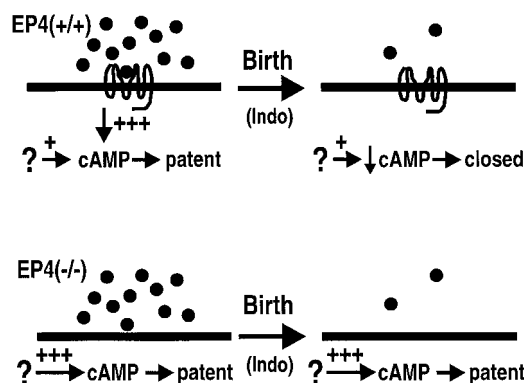


Figure 4 Model for closure of the DA in normal and EP₄-deficient mice. As the fetus approaches term, intracellular cAMP levels in the cells of the DA are maintained largely by the binding of PGE₂ to the EP₄ receptors and stimulation of adenylyl cyclase through G_s. To maintain normal levels of cAMP, other pathways that maintain cAMP levels are downregulated. At birth, a dramatic decrease in circulating levels of PGE₂ occurs. This results in a drop in intracellular cAMP, which in turn triggers the series of events leading to remodelling of the DA. In the absence of the EP₄ receptor, the cells in the DA maintain intracellular cAMP levels independently of PGE₂. As a result, these cells fail to respond to the loss of this pathway at birth. Indomethacin mimics the events that occur at birth by inhibiting synthesis of PGE₂. Black circles, PGE₂; question marks indicate that the sources of cAMP have not been defined. Indo, indomethacin.

DA response to PGE₂ and indomethacin. At E18.5, pregnant mice received a single intravenous dose of indomethacin (10 µg kg⁻¹ in bicarbonate buffer, pH 7.4). At the time indicated, the mice were killed, fetuses removed and immediately placed on ice, and the DA examined. PGE₂ was prepared in ethanol and diluted to 1 µg µl⁻¹ in sterile PBS. Neonates received a single subcutaneous injection of 15 µl; controls received vehicle. Pups were returned to their mother until the time indicated.

Received 17 July; accepted 22 September 1997.

- Heymann, M. A. & Rudolph, A. M. Control of the ductus arteriosus. *Physiol. Rev.* **55**, 62–78 (1975).
- Jager, B. V. & Wollenman, O. J. An anatomical study of the closure of the ductus arteriosus. *Am. J. Pathol.* **18**, 595–613 (1942).
- Challis, J. R. G., Dilley, S. R., Robinson, J. S. & Thorburn, G. D. Prostaglandins in the circulation of the fetal lamb. *Prostaglandins* **11**, 1041–1046 (1976).
- Clyman, R. I., Mauray, F., Roman, C., Rudolph, A. M. & Heymann, M. A. Circulating prostaglandin E₂ concentrations and patent ductus arteriosus in fetal and neonatal lambs. *J. Pediatr.* **97**, 455–461 (1980).
- Clyman, R. I., Wong, L., Heymann, M. A. & Rudolph, A. M. Responsiveness of the lamb ductus arteriosus to prostaglandins and their metabolites. *Prostaglandins* **15**, 325–331 (1980).
- Mitchell, M. D., Brunt, J., Clover, L. & Walker, D. W. Prostaglandins in the umbilical and uterine circulations during late pregnancy in the ewe. *J. Reprod. Fertil.* **58**, 283–287 (1980).
- Piper, P. J., Vane, J. R. & Wyllie, J. H. Inactivation of prostaglandins by the lungs. *Nature* **225**, 600–603 (1970).
- Patterson, D. F. Genetic factors in persistence of the ductus arteriosus. *Paediatr. Cardiol.* **2**, 45–57 (1979).
- Gersony, W. M. Patent ductus arteriosus in the neonate. *Pediatr. Clin. North Am.* **33**, 545–560 (1986).
- Gittenberger-de Groot, A. C., Strengers, J. L. M., Mentink, M., Poelmann, R. E. & Patterson, D. F. Histologic studies on normal and persistent ductus arteriosus in the dog. *J. Am. Coll. Cardiol.* **6**, 394–404 (1985).
- Jarkovska, D., Janatova, T., Hruza, J., Ostadal, B. & Samanek, M. Effect of prostaglandin E₂ on the ductus arteriosus in the newborn rat: An ultrastructural study. *Physiological Reviews* **41**, 323–329 (1992).
- Cocconi, F. & Olley, P. M. The response of the ductus arteriosus to prostaglandins. *Can. J. Physiol. Pharmacol.* **51**, 220–225 (1973).
- Christiansen, N. C. & Fabricius, J. Medical manipulation of the ductus arteriosus. *Lancet* **2**, 24–26 (1975).
- Elliot, R. B., Starling, M. B. & Neutze, J. M. Medical manipulation of the ductus arteriosus. *Lancet* **1**, 140–142 (1975).
- Olley, P. M., Cocconi, F. & Bocach, E. E-type prostaglandins: A new emergency therapy for certain types of congenital heart malformations. *Circulation* **53**, 728–731 (1976).
- Sharpe, G. L., Larsson, K. S. & Thalmé, B. Studies on closure of the ductus arteriosus. *Prostaglandins* **9**, 585–596 (1975).
- Moise, K. J. *et al.* Indomethacin in the treatment of premature labor. Effects on the fetal ductus arteriosus. *N. Engl. J. Med.* **319**, 327–331 (1988).
- Momma, K. & Takao, A. *In vivo* constriction of the ductus arteriosus by nonsteroidal anti-inflammatory drugs in near-term and preterm fetal rats. *Pediatr. Res.* **22**, 567–572 (1987).
- Heymann, M. A., Rudolph, A. M. & Silverman, N. H. Closure of the ductus arteriosus in premature infants by inhibition of prostaglandin synthesis. *N. Engl. J. Med.* **295**, 530–533 (1976).
- Friedman, W. F., Hirschklau, M. J., Printz, M. P., Pitlick, P. T. & Kirkpatrick, S. E. Pharmacologic closure of patent ductus arteriosus in the premature infant. *N. Engl. J. Med.* **295**, 526–529 (1976).
- Coleman, R. A., Smith, W. L. & Narumiya, S. International union of pharmacology classification of prostanoid receptors: properties, distribution, and structure of the receptors and their subtypes. *Pharmacol. Rev.* **46**, 205–229 (1994).

- Smith, G. C. S., Coleman, R. A. & McGrath, J. C. Characterization of dilator prostanoid receptors in the fetal rabbit ductus arteriosus. *J. Pharmacol. Exp. Ther.* **271**, 390–396 (1994).
- Cocconi, F. Control of the ductus arteriosus—a new function for cytochrome P450, endothelin and nitric oxide. *Biochem. Pharmacol.* **48**, 1315–1318 (1994).
- Clyman, R. I. Ductus arteriosus: Current theories of prenatal and postnatal regulation. *Sem. Perinatol.* **11**, 64–71 (1987).
- Cocconi, F. & Olley, P. M. Role of prostaglandins, prostacyclins, and thromboxanes in the control of prenatal patency and postnatal closure of the ductus arteriosus. *Sem. Perinatol.* **4**, 109–113 (1980).
- Honda, A. *et al.* Cloning and expression of a cDNA for mouse prostaglandin E receptor EP₂ subtype. *J. Biol. Chem.* **268**, 7759–7762 (1993).
- Regan, J. W. *et al.* Cloning of a novel human prostaglandin receptor with characteristics of the pharmacologically defined EP₂ subtype. *Mol. Pharmacol.* **46**, 213–220 (1994).
- Engelhardt, J. F., Yankaskas, J. R. & Wilson, J. M. *In vivo* retroviral gene transfer into human bronchial epithelia of xenografts. *J. Clin. Invest.* **90**, 2598–2607 (1992).
- Hooper, M., Hardy, K., Handyside, A., Hunter, S. & Monk, M. HPRT-deficient (Lesch–Nyhan) mouse embryos derived from germline colonization by cultured cells. *Nature* **326**, 292–295 (1987).
- Mohn, A. & Koller, B. H. In *DNA Cloning 4* (eds Glover, D. M. & Hames, B. D.) 143–184 (Oxford University Press, New York, 1995).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper from Mary Sheehan at the London editorial office of Nature.

Acknowledgements. We thank B. Anderson and L. Johnson for assistance in cloning the EP₄ gene; B. Gargies for animal husbandry; A. Latour for help with tissue culture; K. Burns, T. Bartolotta and R. Bagnell for assistance with histology and photomicroscopy; and A. Oakley for help with manuscript preparation. This work was supported by grants from Pfizer, the US NHLBI and the NIDDK.

Correspondence and requests for materials should be addressed to B.H.K. (e-mail: treawouns@aol.com).

Complementation of dominant suppression implicates CD98 in integrin activation

Csilla A. Fenczik, Tariq Sethi*, Joe W. Ramos, Paul E. Hughes & Mark H. Ginsberg

Department of Vascular Biology, The Scripps Research Institute, 10550 North Torrey Pines, La Jolla, California 92037, USA

*Department of Respiratory Medicine, University of Edinburgh, Medical School, Teviot Place, Edinburgh EH8 9AG, UK

The integrin family of adhesion receptors are involved in cell growth, migration and tumour metastasis¹. Integrins are heterodimeric proteins composed of an α and a β subunit, each with a large extracellular, a single transmembrane, and a short cytoplasmic domain. The dynamic regulation of integrin affinity for ligands in response to cellular signals is central to integrin function². This process is energy dependent and is mediated through integrin cytoplasmic domains³. However, the cellular machinery regulating integrin affinity remains poorly understood. Here we describe a genetic strategy to disentangle integrin signalling pathways. Dominant suppression occurs when overexpression of isolated integrin β₁ cytoplasmic domains blocks integrin activation. Proteins involved in integrin signalling were identified by their capacity to complement dominant suppression in an expression cloning scheme. CD98, an early T-cell activation antigen that associates with functional integrins⁴, was found to regulate integrin activation. Furthermore, antibody-mediated crosslinking of CD98 stimulated β₁ integrin-dependent cell adhesion. These data indicate that CD98 is involved in regulating integrin affinity, and validate an unbiased genetic approach to analysing integrin signalling pathways.

Strategies based on genetic selection have proven to be powerful in the dissection of a number of different signal-transduction pathways. We therefore designed a new genetic cloning strategy, based on changes in integrin affinity, to identify potential modulators of integrin activation. This strategy used a Chinese hamster ovary (CHO) cell line (αβpy) stably expressing a chimaeric integrin (α_{IIB}α_{6A}β₃β₁) that contains the extracellular and transmembrane domains of α_{IIB}β₃ fused to cytoplasmic domains of α_{6A}β₁ (ref. 5). This chimaeric integrin has the ligand-binding properties of α_{IIB}β₃ and is activated through the α_{6A}β₁ cytoplasmic domains. In αβpy cells this chimaeric integrin is constitutively active, as measured by

the binding of a ligand mimetic monoclonal antibody, PAC1, which recognizes only the activated form of the chimaeric integrin⁶.

Overexpression of isolated integrin β_1 cytoplasmic domains, in the form of a Tac- β_1 chimaera, results in cell-autonomous, concentration-dependent suppression of integrin signalling^{7,8}. Expression of Tac- β_1 in fibroblast cell lines interferes with cell spreading, migration, fibronectin matrix assembly, and integrin activation^{7,8}. This dominant suppression is structurally specific as neither chimaeric molecules with integrin α -cytoplasmic domains nor certain β -cytoplasmic domain variants have inhibitory activity⁷. Excess free integrin β_1 cytoplasmic domains may cause dominant suppression by titration of essential intracellular proteins that are responsible for

integrin activation (Fig. 1a). We hoped that our strategy allowed us to rescue the dominant suppression caused by Tac- β_1 expression by overexpressing these essential intracellular proteins (Fig. 1a).

To identify proteins that complement Tac- β_1 suppression, $\alpha\beta$ Py cells were co-transfected with a CHO-cell cDNA expression library and Tac- β_1 . Fluorescence-activated cell sorting (FACS) was used to isolate cells in which the chimaeric integrin was in an activated form despite the presence of Tac- β_1 (Fig. 1b). Transfected cDNAs, enriched in this way for complementing activity, were then recovered from the sorted cell populations. Six separate FACS screens resulted in the isolation of 1,200 cDNAs, which were then grouped into 75 pools each of 16 cDNAs. Each pool of cDNAs was screened

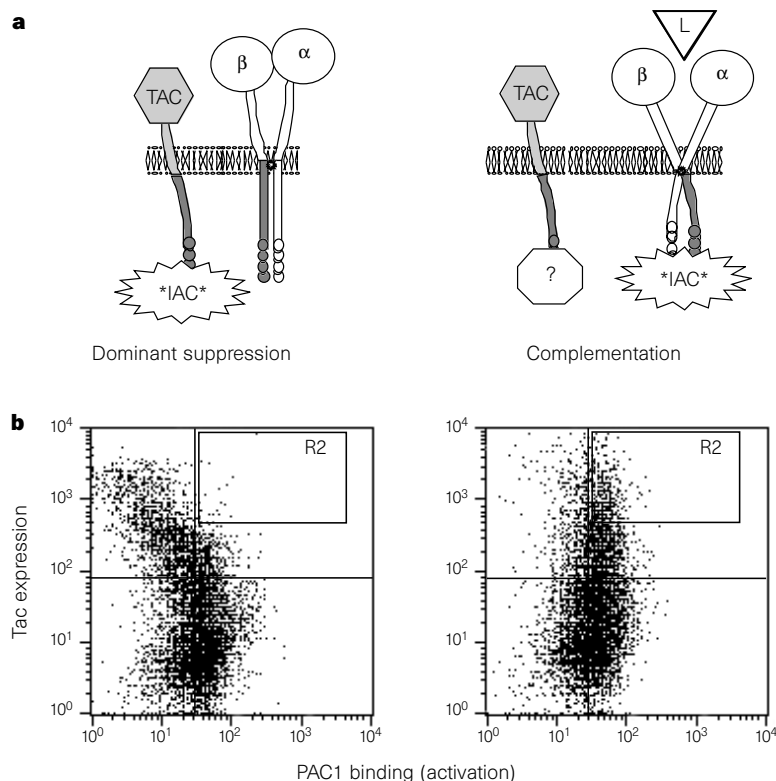


Figure 1 Complementation of dominant suppression and its effect on ligand binding. **a**, Complementation of dominant suppression. Schematic representation of the rationale behind the expression cloning strategy. Intracellular proteins that regulate integrin affinity, shown here as the integrin activation complex (IAC), interact with the cytoplasmic domains of integrins resulting in an increase in ligand binding affinity. Dominant suppression of integrin affinity by Tac- β_1 may be caused by the titration of proteins within the IAC (left). We hypothesized that overexpression of components of IAC, or proteins that block binding to the β cytoplasmic domains of the chimaeric molecule, result in an increase in integrin affinity in the presence of Tac- β_1 (right). **b**, Initial enrichment. $\alpha\beta$ Py cells were transiently transfected with Tac- β_1 and a cDNA expression library. After 48 h, cells were collected and stained for Tac expression (ordinate) and PAC1 binding (abscissa)⁷. Fluorescence-activated cell sorting (FACS) was used to isolate those cells in which the chimaeric integrin is able to bind to PAC1 in the presence of high levels of Tac- β_1 (R2). Shown here are representative dot plots obtained with cells that have been transfected with either Tac- α_5 or Tac- β_1 and analysed by two-colour flow cytometry⁶. A cell in which complementation of dominant suppression has occurred will appear in the upper right-hand quadrant (R2) in the Tac- β_1 transfected cells.

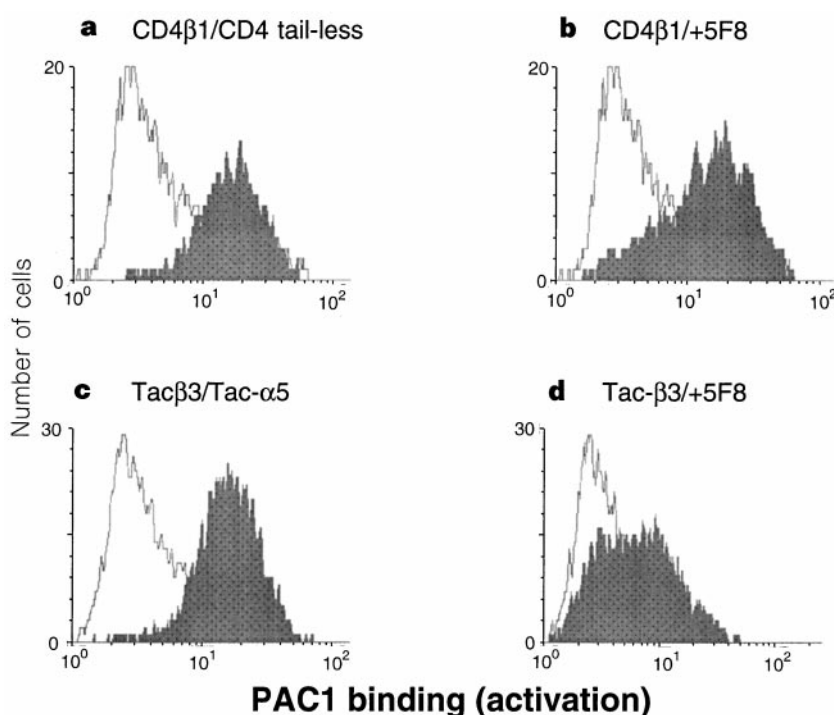


Figure 2 Cytoplasmic domain specificity of complementation of dominant suppression. Data represent PAC1 binding to the CD4- (a, b) or Tac- (c, d) positive subset of transfected $\alpha\beta$ Py cells. **a**, PAC1 binding to $\alpha\beta$ Py cells transfected with CD4- β_1 (ref. 21) (open graph) or with a CD4 construct lacking a cytoplasmic domain (shaded). **b**, Co-transfection with cDNA clone 5F8 and CD4- β_1 is indicated by the shaded graph. PAC1 binding is restored in CD4- β_1 -positive cells co-transfected with 5F8 (shaded). Transfection with 5F8 did not reduce CD4 expression (data not shown). **c**, d, PAC1 binding to $\alpha\beta$ Py cells transfected with Tac chimaeras. Tac- β_3 transfection is indicated by the open graphs. **c**, The shaded graph indicates binding to cells transfected with a non-inhibitory chimaera, Tac- α_5 . **d**, Co-transfection of clone 5F8 with Tac- β_3 restores PAC1 binding (shaded).

for complementing activity by transfecting the pooled cDNAs into $\alpha\beta\text{Py}$ cells along the Tac- β_1 . cDNAs whose expression caused an increase in PAC1 binding even in the presence of high levels of Tac- β_1 were identified as positives. Tac- β_1 expression levels were monitored in all experiments by staining with an anti-Tac antibody (7G7B6) to eliminate those cDNAs with complementation of dominant suppression that resulted from reduced expression of Tac- β_1 . One out of 75 pools contained cDNA that complemented dominant suppression. This pool was divided into four groups of four cDNAs and rescreened as described above, and only one pool

contained a positive cDNA. In this smaller pool, only a single cDNA, 5F8, was able to reverse Tac- β_1 suppression.

The CHO-cell cDNA library was rescreened in four additional experiments that resulted in the enrichment of 1,059 more cDNA clones. Three additional cDNAs identical to 5F8 were isolated from these clones. Consequently, in this cDNA library, 5F8 was unique in complementing suppression. This library has a reported minimal complexity of 1.8×10^7 independent clones. Thus the effect of 5F8 appears to be highly specific. Transfection of cDNAs reported to interact with or regulate integrins, including α -actinin, vinculin, paxillin, integrin-linked kinase, and β_3 -endoneurin, had no effect on Tac- β_1 suppression of integrin activation (data not shown). This result further supports the idea that 5F8 complementation of dominant suppression is very specific.

To determine if the rescue by 5F8 is dependent on the cytoplasmic or extracellular domain of Tac- β_1 , we assessed the ability of 5F8 to complement suppression initiated by a β_1 tail and transmembrane domain joined to the extracellular domain of CD4 (Fig. 2). As with Tac- β_1 , co-transfection of 5F8 with CD4- β_1 resulted in a reversal of dominant suppression, indicating that the rescue was independent of the extracellular domain of the suppressive chimera. The isolated β_3 cytoplasmic domain is similar to the β_1 cytoplasmic domain and can also initiate dominant suppression^{7,8}. We therefore investigated whether rescue by 5F8 is specific to the β_1 cytoplasmic domain. We found that 5F8 also reversed Tac- β_3 suppression (Fig. 2). Thus 5F8 can complement suppression initiated by either β_1 or β_3 cytoplasmic domains.

The 5F8 sequence has been deposited in Genbank (accession no. U93712) and contains 1,902 base pairs, with an open reading frame encoding a polypeptide of 533 amino acids. In addition, it contains a poly(A) tract and polyadenylation signal. Analysis of the predicted topology of the encoded protein using Tmpred⁹ suggested that it possesses at least one transmembrane domain with the amino terminus inside the cell (type II transmembrane protein). Furthermore, a BLAST¹⁰ database search showed that the amino-acid sequence encoded by 5F8 is 72% identical to that of the heavy chain of the 4F2 antigen, CD98 (ref. 11). Overexpression of human CD98 also reversed dominant suppression, and so we conclude that we have isolated its hamster homologue (Fig. 3a).

CD98 was originally identified as an early T-cell activation antigen. It is part of a heterodimer of relative molecular mass 120,000 (M_r 120K), consisting of a heavy chain of approximately 80K that bears CD98 epitopes, and a light chain of 40K (ref. 12). It is a membrane protein that specifically associates with β_1 integrins⁴. Indeed, it has been shown that CD98 co-precipitated with functional $\alpha_2\beta_1$, $\alpha_3\beta_1$, $\alpha_5\beta_1$ and $\alpha_6\beta_1$ integrins⁴. Furthermore, co-precipitation with $\alpha_4\beta_1$ was only observed when $\alpha_4\beta_1$ was activated in the presence of 2 mM Mn^{2+} . The light chain is poorly characterized, and its cDNA has yet to be cloned. Thus we cannot assess the role of the light chain in the effect of CD98 on integrin activation.

To determine whether the rescue of dominant suppression is a function of other integrin-associated membrane proteins, we overexpressed membrane proteins previously implicated in integrin function: CD9 (ref. 13), CD47 (integrin-associated protein)¹⁴, and urokinase plasminogen activator receptor (uPAR)¹⁵; in each case, protein expression was confirmed by flow cytometry (data not shown). Expression of these proteins failed to reverse dominant suppression by Tac- β_1 (Fig. 3a). Furthermore, overexpression of CD98 or the other membrane-associated proteins in the absence of Tac- β_1 did not increase the affinity of already activated integrins (Fig. 3b). Thus the rescue of Tac- β_1 suppression by CD98 is not an artefact of overexpression of membrane proteins.

The $\alpha_3\beta_1$ integrin-mediated adhesion of the small cell lung-cancer cell line (SCLC) H345 to laminin and fibronectin can be regulated dynamically¹⁶. We therefore tested whether CD98 might be involved in the regulation of integrins in SCLC adhesion. An anti-CD98 monoclonal antibody (4F2) markedly enhanced single-

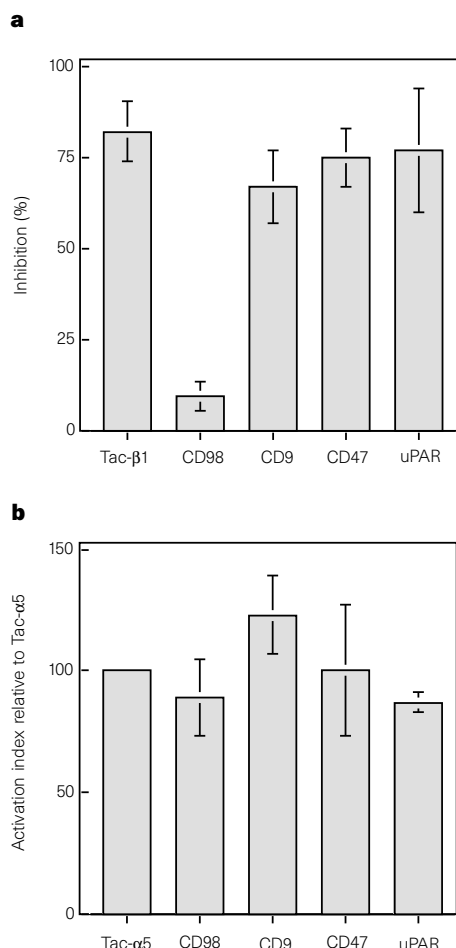


Figure 3 Human CD98 complements dominant suppression, whereas several other membrane proteins do not. $\alpha\beta\text{Py}$ cells were transfected with 2 μg of either Tac- β_1 (a) or Tac- α_5 (b). They were simultaneously transfected with 4 μg of cDNAs encoding either human CD98 (ref. 11), CD9 (ref. 22), CD47 (ref. 14) or uPAR. After 48 h, cells were collected and analysed for PAC1 binding to the Tac-positive subset of cells. To obtain quantitative estimates of integrin activation we calculated a numerical activation index defined as $100 (F_o - F_R) / (F_{\text{LIBS6}} - F_R)$, as previously described²⁵, where F_o is median fluorescence intensity of PAC1 binding; F_R is the background fluorescence intensity of PAC1 binding in the presence of a competitive inhibitor (1 μM Ro43-5054), and F_{LIBS6} is the maximal fluorescence intensity in the presence of 2 μM anti-LIBS6, an activating monoclonal antibody. The mean $\pm$ s.d. of three independent experiments for each membrane protein is shown. **a**, Specificity of CD98 complementation. Percentage inhibition was defined as $100(A_0 - A)/A_0$, where A_0 is the activation index in Tac- α_5 transfected cells. A is the activation index in Tac- β_1 transfected cells. The co-transfected membrane protein is indicated below each column. **b**, Effect of membrane proteins in the absence of dominant suppression. Cells were transfected with Tac- α_5 alone or the indicated membrane protein. Data are expressed as relative activation = $100(A_{\text{MP}}/A_{\alpha_5})$, where A_{MP} is the activation index in the presence of the membrane protein and A_{α_5} is the activation index with Tac- α_5 alone. The activation index with Tac- α_5 alone was 60%.

cell adhesion of H345 cells to both laminin and fibronectin (Fig. 4). Addition of the function blocking anti- β_1 antibody (P5D2) or 2 mM EDTA abrogated this effect. Control antibodies for 4F2 (D57, anti- $\alpha_{IIb}\beta_3$) and for P5D2 (P41, anti- β_3) had no effect. In addition, B6H12, a monoclonal antibody against CD47 (ref. 17), did not enhance adhesion to fibronectin or laminin (results not shown). To assess the role of antibody-mediated CD98 crosslinking in enhancing integrin function, we examined the effects of monovalent Fab fragments of 4F2 (Fab-4F2) on SCLC cell adhesion. Fab-4F2 did not enhance SCLC cell adhesion to laminin or fibronectin (Fig. 4). The binding of the Fab fragments was confirmed by their ability to block the enhanced adhesion caused by the intact antibody. Consequently, clustering of CD98 is required for 4F2-induced upregulation of β_1 integrin function.

Our genetic complementation strategy has led to the identification of CD98 as a protein involved in integrin activation. Indeed, CD98 specifically associates with β_1 integrins, and that association may be activation dependent, as CD98 is preferentially immunoprecipitated with activated integrins⁴. In addition, CD98 lacking a cytoplasmic domain fails to complement dominant suppression (C.A.F., B. Lewis & M.H.G., unpublished results). This result indicates that the association between integrins and CD98 may be mediated by the cytoplasmic domains of both proteins. It is likely that endogenous hamster CD98 is expressed in CHO cells, as 5F8 was isolated from a cDNA library derived from CHO-cell mRNA and CD98 is expressed in many tissue-culture cell lines¹¹. Cross-linking CD98 with 4F2 antibody stimulated β_1 -dependent adhe-

sion. Thus the rescue of dominant suppression by overexpression of CD98 in $\alpha\beta$ py cells may work because overexpression of CD98 may favour clustering. Further work will be required to establish relationships between CD98 clustering, its association with integrins, and the identity of CD98 ligands that might cluster CD98.

Our finding that CD98 is involved in integrin function explains the observation that CD98-mediated membrane fusion of cells transfected with HIV gp160 is integrin dependent^{18,19}. Further, when T cells are activated, CD98 expression increases before an increase in β_1 integrin-mediated adhesion²⁰. Therefore changes in integrin function in these two cell types could depend upon CD98 regulation. Collectively, these data indicate that CD98 may be an important regulator of integrin function in many contexts. □

Methods

cDNA constructs and antibodies. Expression plasmids encoding the extracellular domain of murine CD4 fused to the human β_1 transmembrane and cytoplasmic domains and CD4 tailless, which lacks the β_1 cytoplasmic domain, were a gift from M. Lukashev²¹. Human 4F2 antigen (CD98) cDNA was provided by J. M. Leiden¹¹, and was subcloned into pCDNA1 as an *Eco*R1 fragment. The pCDM8 construct encoding CD9 has been described²². The pRc/RSV plasmid, pIAP45, encoding CD47, and B6H12, the anti-CD47 antibody, were gifts from F. Lindberg and E. Brown¹⁴. cDNA encoding uPAR, in pCDNA1, pcuPAR1, was provided by L. Miles. The hybridoma cell line 4F2(C13) was purchased from American Type Culture Collection (ATCC). Ascites were produced in pristane-primed BALB/c mice. Fab fragments were prepared by papain digestion of purified 4F2 IgG (2 mg ml⁻¹) for 5 h at 37 °C. Digestion was terminated by the addition of iodoacetamide. Fab fragments were purified on Protein-A sepharose columns. Fab fragments were characterized by SDS-PAGE and exhibited characteristic mobilities.

Expression cloning. A cDNA library, made from poly(A)⁺ mRNA from CHO-K1 cells, directionally cloned into the *Not*I site of the mammalian expression vector, pCDNA1, was purchased from Invitrogen (San Diego, CA). The library is reported to contain 1.8×10^7 primary recombinants. A CHO cell line that expresses the polyoma large T antigen and an activated recombinant chimaeric $\alpha_{IIb}\beta_3$ integrin ($\alpha\beta$ py cells) was constructed as described⁵. $\alpha\beta$ py cells were transfected with the Tac- β_1 (Tac- β_1 , Tac- β_3 and the non-inhibitory Tac- α_5 chimaeras were gifts from S. LaFlamme and K. Yamada⁸; 4 μ g per plate) using lipofectamine (Gibco, BRL), according to the manufacturer's instructions. The transfection efficiency ranged from 30% to 50%, as determined by the binding of the anti-Tac antibody (7G7B6)⁷. After 48 h incubation, cells were collected and stained as described⁶, and cells positive for both PAC1 and Tac were isolated by fluorescence-activated cell sorting (FACSTAR, Becton Dickinson). Plasmid DNA was extracted²³ and digested with *Dpn*I to remove plasmids that were not replicated in transfected cells, and used to transform *Escherichia coli* MC1061/P3. Individual colonies were grown, and the bacteria were then pooled into groups of 16 for plasmid purification (Qiagen, Chatsworth, CA). Groups of cDNAs were transfected into the $\alpha\beta$ py cell line along with Tac- β_1 and the transfectants were screened by flow cytometry as described above. Pools containing cDNAs that altered PAC1 binding were further screened in smaller pools of four, and then the positive pools were screened as individual colonies.

Adhesion assay. Laminin and fibronectin (Calbiochem) were used as substrates in serum-free adhesion assays. Cells of the small cell lung-cancer cell line H345 (ATCC) were maintained in RPMI and 10% serum and for experimental purposes were passaged into serum-free medium RPMI 1640 containing SITA (Sigma). Cells ($1-2 \times 10^6$ per ml) were washed twice in RPMI 3.5 d post-passage and disaggregated into single cells. Cells in RPMI (50 μ l) were added to 96-well tissue-culture plates (Costar) coated with extracellular matrix proteins blocked with 1 mg ml⁻¹ BSA. Cells were allowed to attach for 45 min at 37 °C. Cell attachment was determined by crystal violet staining. The attachment of H345 cells to wells coated with 25 mg ml⁻¹ poly-L-lysine and fixed with 20% glutaraldehyde before aspiration was defined as 100% adhesion²⁴.

Received 13 June; accepted 8 September 1997.

- Hynes, R. O. Integrins: versatility, modulation and signaling in cell adhesion. *Cell* **69**, 11–25 (1992).
- Schwartz, M. A., Schaller, M. D. & Ginsberg, M. H. Integrins: emerging paradigms of signal transduction. *Annu. Rev. Cell Biol.* **11**, 549–599 (1995).
- Williams, M. J., Hughes, P. E., O'Toole, T. E. & Ginsberg, M. H. The inner world of cell adhesion: integrin cytoplasmic domains. *Trends Cell Biol.* **4**, 109–112 (1994).

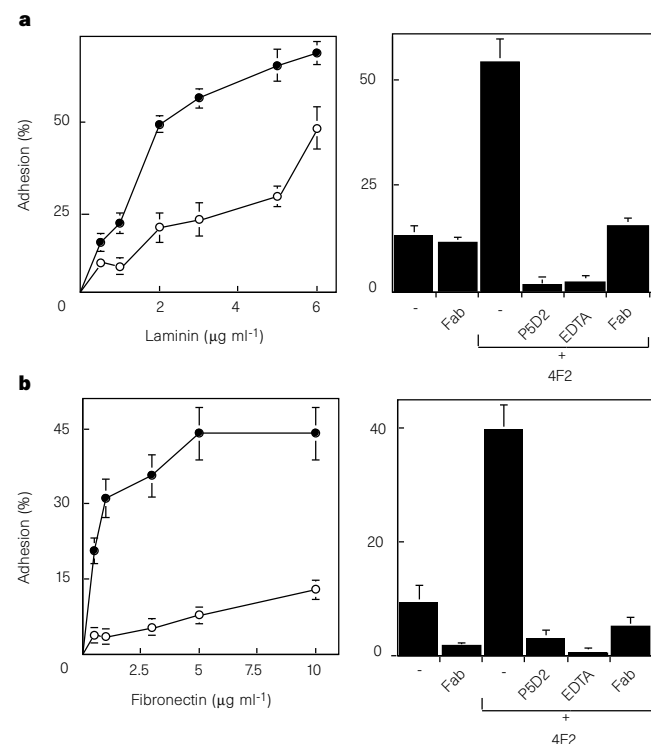


Figure 4 The effect of a 4F2 monoclonal antibody on adhesion of the small cell lung-cancer cell line (SCLC) H345 to extracellular matrix. Left, attachment of SCLC cells to 96-well tissue-culture plates coated with increasing concentrations of laminin (**a**) or fibronectin (**b**) in the presence (filled circles) or absence (open circles) of 20 μ g ml⁻¹ 4F2 monoclonal antibody. Right, addition of 4F2-Fab (100 μ g ml⁻¹) to SCLC cells does not increase adhesion to plates coated with 3 μ g ml⁻¹ laminin (**a**) or 10 μ g ml⁻¹ fibronectin (**b**); addition of P5D2 (a β_1 -integrin function-blocking antibody) or 2 mM EDTA also block the increase of adhesion of SCLC cells to plates coated with either 3 μ g ml⁻¹ laminin (**a**) or 10 μ g ml⁻¹ fibronectin (**b**). Results show the mean percentage adhesion above background (which was consistently <5%) compared with poly-L-lysine (taken as 100%) of 4–6 independent experiments in duplicate/triplicate – s.e.m.

4. Mannion, B. A., Berditchevski, F., Eble, J. A. & Hemler, M. E. Functionally active β_1 integrins show specific association with CD98 protein. (submitted).
5. Baker, E. K. *et al.* A genetic analysis of integrin function: Glanzmann thrombasthenia in vitro. *Proc. Natl Acad. Sci. USA* **94**, 1973–1978 (1997).
6. O'Toole, T. E. *et al.* Integrin cytoplasmic domains mediate inside-out signal transduction. *J. Cell Biol.* **124**, 1047–1059 (1994).
7. Chen, Y. *et al.* "Inside-Out" signal transduction inhibited by isolated integrin cytoplasmic domains. *J. Biol. Chem.* **269**, 18307–18310 (1994).
8. LaFlamme, S. E., Thomas, L. A., Yamada, S. S. & Yamada, K. M. Single subunit chimeric integrins as mimics and inhibitors of endogenous integrin functions in receptor localization, cell spreading and migration, and matrix assembly. *J. Cell Biol.* **126**, 1287–1298 (1994).
9. Hofmann, K. & Stoffel, W. F. A database of membrane spanning protein segments. *Hoppe-Seyler's Z. Biol. Chem.* **347**, 166–180 (1993).
10. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
11. Quackenbush, E. *et al.* Molecular cloning of complementary DNAs encoding the heavy chain of the human 4F2 cell surface antigen: a type II membrane glycoprotein involved in normal and neoplastic cell growth. *Proc. Natl Acad. Sci. USA* **84**, 6526–6530 (1987).
12. Haynes, B. F. *et al.* Characterization of a monoclonal antibody (4F2) that binds to human monocytes and to a subset of activated lymphocytes. *J. Immunol.* **126**, 1409–1414 (1981).
13. Berditchevski, F., Zutter, M. M. & Hemler, M. E. Characterization of novel complexes on the cell surface between integrins and proteins with 4 transmembrane domains (TM4 proteins). *Mol. Biol. Cell* **7**, 193–207 (1996).
14. Lindberg, F. P., Gresham, H. D., Schwarz, E. & Brown, E. J. Molecular cloning of integrin-associated protein: an immunoglobulin family member with multiple membrane spanning domains implicated in $\alpha_5\beta_1$ -dependent ligand binding. *J. Cell Biol.* **123**, 485–496 (1993).
15. Wei, Y. *et al.* Regulation of integrin function by the urokinase receptor. *Science* **273**, 1551–1555 (1996).
16. Elices, M. J., Urry, L. A. & Hemler, M. E. Receptor functions for the integrin VLA-3: Fibronectin, collagen and laminin binding are differentially influenced by ARG-GLY-ASP peptide and by divalent cations. *J. Cell Biol.* **112**, 169–181 (1991).
17. Brown, E., Hooper, L., Ho, T. & Gresham, H. Integrin-associated protein: a 50-kD plasma membrane antigen physically and functionally associated with integrins. *J. Cell Biol.* **111**, 2785–2794 (1990).
18. Ohgimoto, S. *et al.* Molecular characterization of fusion regulatory protein-1 (FRP-1) that induces multinucleated giant cell formation of monocytes and HIV gp160-mediated cell fusion. FRP-1 and 4F2/CD98 are identical molecules. *J. Immunol.* **155**, 3585–3592 (1995).
19. Ohta, H. *et al.* Molecular and biological characterization of fusion regulatory proteins (FRPs): anti-FRP mAbs induced HIV-mediated cell fusion via an integrin system. *EMBO J.* **13**, 2044–2055 (1994).
20. Shimizu, Y., Van Seventer, G. A., Horgan, K. J. & Shaw, S. Regulated expression and binding of three VLA (β_1) integrin receptors on T cells. *Nature* **345**, 250–253 (1990).
21. Lukashev, M. E., Sheppard, D. & Pytela, R. Disruption of integrin function and induction of tyrosine phosphorylation by the autonomously expressed β_1 integrin cytoplasmic domain. *J. Biol. Chem.* **269**, 18311–18314 (1994).
22. Indig, F. E., Diaz-Gonzalez, F. & Ginsberg, M. H. Analysis of the tetraspanin CD9-integrin $\alpha_{IIb}\beta_3$ (GPIIb-IIIa) complex in platelet membranes and transfected cells. *Biochem. J.* **326** (in the press).
23. Hirt, B. Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.* **26**, 365–369 (1967).
24. Hirawasa, M., Shijubo, N., Ueda, T. Y. & Abe, S. Integrin expression and ability to adhere to extracellular matrix proteins and endothelial cells in human lung cancer lines. *Br. J. Cancer* **70**, 466–473 (1994).
25. Hughes, P. *et al.* Suppression of integrin activation: A novel function of a Ras/Raf-initiated MAP kinase pathway. *Cell* **88**, 521–530 (1997).

Acknowledgements. We thank L. Baker for generation of the $\alpha\beta$ Py cell line and for discussions; S. Shattil and M. Schwartz for suggestions, and reviews of the manuscript; and our colleagues for the reagents acknowledged in the text. This work was supported by grants from the NIH and Cor Therapeutics. T.S. is an MRC (UK) travelling fellow, and P.E.H. is a fellow of the Leukemia Society of America.

Correspondence and requests for materials should be addressed to M.H.G. (e-mail: ginsberg@scripps.edu)

Inhibitory and activating functions for MAPK Kss1 in the *S. cerevisiae* filamentous-growth signalling pathway

Jeanette Gowen Cook*†, Lee Bardwell* & Jeremy Thorner

Department of Molecular and Cell Biology, Division of Biochemistry and Molecular Biology, University of California, Berkeley, California 94720-3202, USA

* These authors contributed equally to this work.

Mitogen-activated protein kinase (MAPK) cascades are conserved signalling modules that regulate responses to diverse extracellular stimuli, developmental cues and environmental stresses (reviewed in refs 1–3). A MAPK is phosphorylated and activated by a MAPK kinase (MAPKK), which is activated by an upstream protein kinase, such as Raf, Mos or a MAPKK kinase. Ste7, a MAPKK in the yeast *Saccharomyces cerevisiae*, is required for two developmental pathways: mating⁴ and invasive (filamentous) growth^{5,6}. Kss1 and Fus3, the MAPK targets of Ste7, are required for mating^{7,8}, but their role in invasive growth has been unclear.

Because no other *S. cerevisiae* MAPK has been shown to function in invasive growth, it was proposed^{5,6,9} that Ste7 may have non-MAPK targets. We show instead that Kss1 is the principal target of Ste7 in the invasive-growth response in both haploids and diploids. We demonstrate further that Kss1 in its inactive form is a potent negative regulator of invasive growth. Ste7 acts to relieve this negative regulation by switching Kss1 from an inhibitor to an activator. These results indicate that this MAPK has a physiologically important function in its unactivated state. Comparison of normal and MAPK-deficient cells indicates that nitrogen starvation and activated Ras stimulate filamentous growth through both MAPK-independent and MAPK-dependent means.

To examine the role of Ste7, Kss1 and Fus3 in invasive growth, we analysed a set of eight haploid strains in the Σ 1278b background, because this lineage readily displays invasive growth when stimulated by the appropriate medium^{5,6,10}. These strains represent all possible combinations of wild-type and null (deletion) alleles of *STE7*, *KSS1* and *FUS3*, but are otherwise isogenic (Fig. 1a). All of the strains grew equally vigorously (Fig. 1b). As expected, the only strains able to mate were those containing an intact MAPK cascade: JCY100 (*STE7⁺ KSS1⁺ FUS3⁺*), JCY110 (*STE7⁺ kss1 Δ FUS3⁺*), and JCY120 (*STE7⁺ KSS1⁺ fus3 Δ*) (Fig. 1c).

The standard agar penetration assay was used to score for invasive growth (Fig. 1d). We found, in agreement with a previous report⁵, that a strain lacking *KSS1* (JCY110) was unable to invade agar medium, whereas a strain (JCY120) lacking *FUS3* appeared to invade better than wild-type cells (JCY100). As also reported previously⁵, a strain (JCY107) lacking *STE7* or a strain (JCY127) lacking both *STE7* and *FUS3* were unable to invade agar, whereas a strain (JCY130) lacking both *KSS1* and *FUS3* invaded at least as well as wild-type cells. These data had been interpreted⁵ as suggesting that some Ste7-dependent (but MAPK-independent) process promotes invasive growth. To test this notion directly, we examined a *ste7 Δ kss1 Δ fus3 Δ* triple-mutant strain (JCY137), which had not previously been constructed. Remarkably, this triple mutant invaded agar medium just as well as its wild-type parent. Hence, the inability of a Ste7-deficient strain to invade agar medium is completely reversed when both Kss1 and Fus3 are also removed (Fig. 1d; compare strains JCY107 and JCY137), indicating that the MAPKs, especially Kss1, act to repress haploid invasive growth, and that the principal function of Ste7 is to overcome this repression. Indeed, by examining a *ste7 Δ kss1 Δ* double mutant, which had also not been tested previously, we found that the inability of a Ste7-deficient cell to invade agar could be suppressed by removing just *KSS1* (Fig. 1d; compare JCY117 and JCY107).

Transcription from filamentous response elements (FREs), which contain the binding sites for two transcription factors (Ste12 and Tec1) required for invasive growth^{5,6,11}, has been shown to correlate with Ras- and MAPK cascade-dependent signalling in this pathway^{12,13}. Quantitative measurement of the expression of an *FRE-lacZ* reporter gene in our set of eight isogenic strains (Fig. 1e) confirmed the results obtained using the agar penetration assay (Fig. 1d). For example, compared with the *ste7 Δ* single mutant (JCY107), full FRE-dependent transcription was restored in the *ste7 Δ kss1 Δ fus3 Δ* triple mutant (JCY137). Deletion of *STE12* in the *ste7 Δ kss1 Δ fus3 Δ* background (YLB637) eliminated both FRE-dependent transcription (Fig. 1e) and abolished invasive growth (data not shown), demonstrating that the MAPK-independent activity of Ste12 can promote FRE-dependent expression.

To determine whether Ste11, the MAPKK kinase that phosphorylates and activates Ste7 (ref. 14), contributes to the invasive growth observed in the *ste7 Δ kss1 Δ fus3 Δ* triple mutant, we generated a *ste11 Δ ste7 Δ kss1 Δ fus3 Δ* quadruple mutant. We found that both *FRE-lacZ* expression and invasive growth were undiminished (data not shown). Thus the entire MAPK cascade previously thought to be necessary for invasive growth is, in a formal sense, dispensable for this developmental pathway.

† Department of Genetics, Duke University Medical Center, Box 3054, Durham, North Carolina 27710, USA.

To confirm that Kss1 functions as an inhibitor of invasive growth and that Ste7 serves to alleviate the negative effect of Kss1, we reintroduced the *KSS1* gene, the *STE7* gene, or both, under the control of their endogenous promoters on *CEN* (low-copy) plasmids, into the triple mutant (JCY137) (Fig. 2a). As expected, only when a functional MAPK cascade (Ste7 plus Kss1) was reconstituted was the ability to mate restored (Fig. 2b). A very different pattern emerged when invasive growth was scored. When *KSS1* alone was introduced (in the absence of *STE7*), agar penetration was potently inhibited (Fig. 2c); *STE7* alone, by contrast, had no effect. The inhibitory effect of *KSS1* was relieved completely, however, when *STE7* was co-introduced with *KSS1* (Fig. 2c). These findings corroborate the conclusion that *STE7* function is dispensable for haploid invasive growth if *KSS1* is absent, but not if *KSS1* is present.

In the absence of Ste7, *KSS1* acts as a strong repressor of agar penetration (Figs 1d, 2c) and FRET-mediated transcription (Fig. 1e), which is readily observed by comparing the properties of the *ste7Δkss1Δfus3Δ* triple-mutant strain (JCY137) to its otherwise isogenic *KSS1*⁺ counterpart (JCY127). In *STE7*⁺ cells, however, *KSS1* also plays a positive role in haploid invasive growth⁵, as deletion of *KSS1* in an otherwise wild-type background (JCY110) resulted in greatly reduced agar penetration (Fig. 1d) and in detectable reduction of FRET-mediated transcription (Fig. 1e) (see also ref. 12). These phenotypes could be complemented by the wild-type *KSS1* gene, but not by a catalytically inactive allele, *kss1*(K42R Q45P)⁸, or by an unactivatable allele, *kss1*(T183A Y185F)⁸ (Fig. 3). In contrast, both the catalytically inactive and unactivatable derivatives, as well as functional Kss1, were all able to repress haploid invasive growth in the *ste7Δkss1Δfus3Δ* triple mutant (Fig. 3). Thus the positive and negative functions of Kss1 in invasive growth are separable; the stimulatory function requires the catalytic activity of Kss1 and its activation by Ste7, whereas the inhibitory function of

Kss1 does not. Signal-dependent conversion of Kss1 from an inhibitor into an activator of invasive growth provides a mechanism for a switch-like transition to initiate this developmental pathway.

In *MATa*/*MATα* diploid cells, *STE7* is required for pseudohyphal development⁶, whereas no role has been attributed to *KSS1* or *FUS3* in this filamentous growth process. Indeed, *FUS3* is not expressed in diploids, even under conditions of nitrogen deprivation, which induce filament formation⁶, whereas *KSS1* is expressed in both diploid and haploid cells¹⁵. We reasoned that if the main role of Ste7 in pseudohyphal development is to relieve Kss1-mediated repression, then a *ste7Δ/ste7Δkss1Δ/kss1Δ* homozygous diploid should be capable of forming filaments. Indeed, such a strain (JCY119) readily formed pseudohyphae on low-nitrogen plates (Fig. 4a), as judged by the standard visual inspection of colony morphology^{6,10}. To examine the contribution of individual genes, either *KSS1* or *STE7*, or both, were reintroduced on *CEN* plasmids into the *ste7Δ/ste7Δkss1Δ/kss1Δ* diploid strain. The presence of *KSS1* (Fig. 4b), but not *STE7* (Fig. 4c), completely inhibited filament formation; correspondingly, in this same strain background, a *ste7Δ/ste7ΔKSS1*^{+/+}/*KSS1*⁺ diploid is deficient in pseudohyphal development (data not shown). Most strikingly, pseudohyphal growth was restored to *KSS1*-expressing cells when *STE7* was also co-introduced (Fig. 4d). Thus, as we demonstrated for haploid invasive growth, *STE7* function is required for diploid pseudohyphal development only if *KSS1* is also present. As observed for wild-type diploids¹⁰, induction of pseudohyphal development in the *ste7Δ/ste7Δkss1Δ/kss1Δ* diploid strain only occurred on low-nitrogen medium (data not shown), indicating that absence of the MAPK cascade does not bypass the requirement for a nutrient limitation signal.

RAS2^{G19V}, the yeast homologue of the activated *ras* oncogene, stimulates filament formation^{10,12}. When introduced on a plasmid

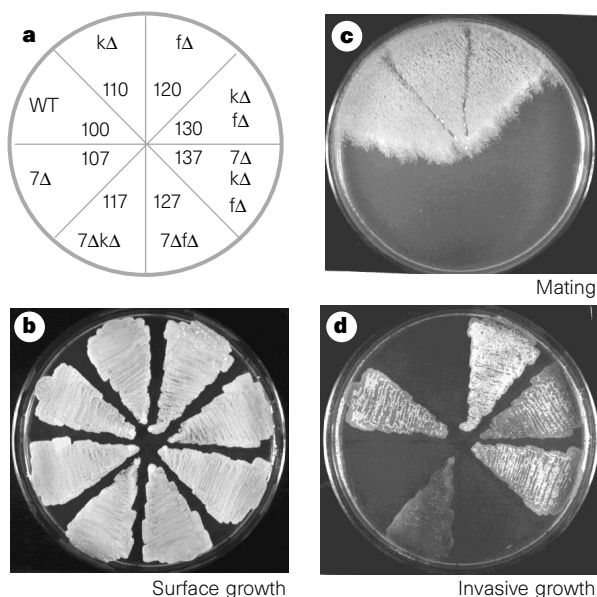
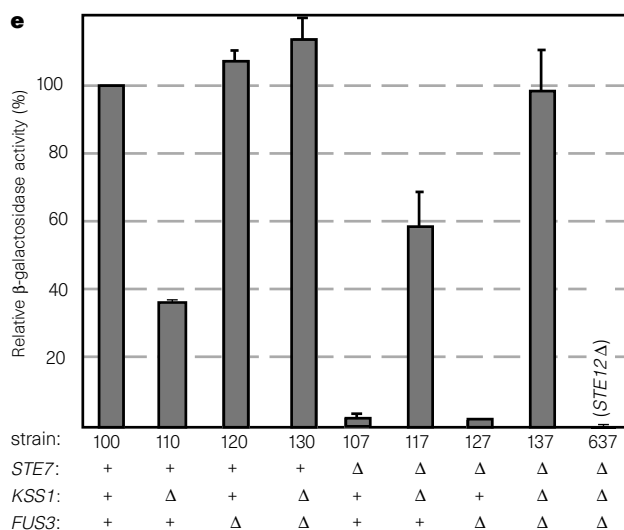


Figure 1 The Ste7 MAPK is not required for haploid invasive growth. **a**, Genotypes of the strains examined. Abbreviations are: WT, wild type; 7Δ, *ste7Δ*; kΔ, *kss1Δ*; and fΔ, *fus3Δ*. **b**, Total growth. Σ 1278b-derived strains, JCY100 (*MATa KSS1*⁺ *FUS3*⁺ *STE7*⁺) and its otherwise isogenic derivatives, JCY110 (*kss1Δ*), JCY120 (*fus3Δ*), JCY130 (*kss1Δ fus3Δ*), JCY107 (*ste7Δ*), JCY117 (*ste7Δ kss1Δ*), JCY127 (*ste7Δ fus3Δ*) and JCY137 (*ste7Δ kss1Δ fus3Δ*) were streaked onto YPD plates¹⁶, grown for 3 days at 30 °C, and photographed. **c**, Mating. The plate shown in **b** was replica-plated onto a lawn of mating tester strain, DC17 (*MATα his1*), allowed to mate for 8 h, and then replica-plated onto medium selective for diploids. **d**, Haploid invasive growth. The plate shown in **b** was rinsed under a stream of running water to distinguish surface growth from invasive growth by the degree of agar penetration⁵. When the kinetics of agar penetration was examined



by incubating replicate plates for different lengths of time before rinsing (data not shown), the invasion-competent strains displayed the following relative efficacies: JCY120 > JCY100 = JCY130 = JCY137 > JCY117. **e**, Expression of the *FRE*^{T1}-*lacZ* reporter gene. The eight strains described in **a** plus another otherwise isogenic strain, YLB637 (*ste7Δkss1Δfus3Δste12Δ*), were transformed with plasmid YEpU-FTyZ, and β -galactosidase activity was measured. Values are normalized to that observed for the wild-type control and represent the averages of measurements, made in duplicate, on protein extracts prepared from three independent transformants of each strain (error bars indicate standard deviation). The specific activity of β -galactosidase expressed in the control strain (JCY100) was 4,280 nmol per min per mg protein.

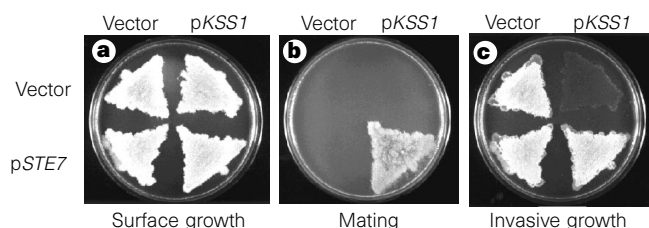


Figure 2 Kss1 is a potent repressor of haploid invasive growth. Strain JCY137 (*MATa ste7Δ kss1Δ fus3Δ*) was transformed with either two 'empty' vectors, YCpT and YCpL (top left), or YCpT and YCpL-KSS1 (top right), or YCpT-STE7 and YCpL (bottom left), or YCpT-STE7 and YCpL-KSS1 (bottom right). The resulting transformants were streaked onto plates selective for the maintenance of both plasmids. After 2 days at 30°C, plates were replica-plated onto YPD plates¹⁸, and scored for total growth (a), mating (b), and invasive growth after 30 h, at 30°C (c).

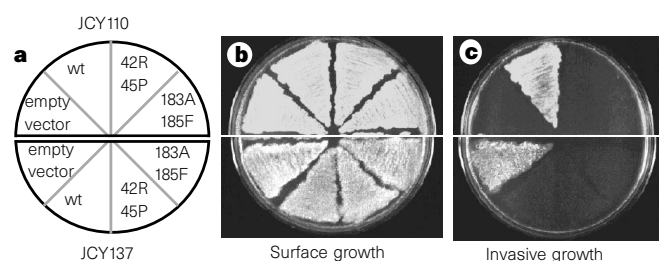


Figure 3 Stimulatory and inhibitory functions of Kss1 are separable. **a**, Plasmids YCpU, YCpU-KSS1, YCpU-KSS1(*K42R Q45P*) and YCpU-KSS1(*T183A Y185F*) were introduced by transformation into strain JCY110 (*MATa kss1Δ*) to assess the ability to provide the positive (stimulatory) function of KSS1, or into strain JCY137 (*MATa ste7Δ kss1Δ fus3Δ*) to assess the negative (inhibitory) function of KSS1. The resulting transformants were streaked onto plates selective for plasmid maintenance. After 2 days at 30°C, plates were replica-plated onto YPD plates¹⁸, and scored for total growth (b) and invasive growth (c) after 24 h (JCY137) or 30 h (JCY110) at 30°C. As shown before⁹, only the wild-type KSS1 allele can complement the mating defect of a *kss1Δ fus3Δ* strain, yet all of the encoded proteins are equally stable. Ability (or lack of ability) of these *kss1* alleles to be phosphorylated *in vivo*⁹ and *in vitro*²⁰, and their protein kinase activity (or lack thereof)²⁰, have been documented in detail elsewhere.

into the *ste7Δ/ste7Δkss1Δ/kss1Δ* diploid strain, this allele stimulated even more florid filament formation (Fig. 4e). Thus the mechanism by which activated Ras stimulates pseudohyphal development does not strictly require the MAPK cascade. When KSS1 was co-introduced, the Ras-stimulated filament formation was detectably and reproducibly reduced (Fig. 4f), in keeping with the negative regulatory role of Kss1 we have demonstrated here. However, we noted that when both STE7 and KSS1 were introduced together, there was a significant and reproducible potentiation of Ras-stimulated filament formation (Fig. 4g), suggesting that the Ras signal is propagated through both MAPK cascade-independent and MAPK cascade-dependent mechanisms acting in parallel (Fig. 5).

Our findings demonstrate that the primary function of the Ste11–Ste7–Kss1 cascade is to repress the initiation of haploid invasive growth and diploid pseudohyphal development. Repression is mediated by Kss1 in its unactivated state, and activation of this MAPK by Ste7 acts to relieve its repressive action (Fig. 5). Activated Kss1 has a positive function in both haploid and diploid invasiveness. In haploids, this role is, in effect, to override the negative function of Fus3, as a *kss1Δ fus3Δ* double mutant displays invasive growth, whereas a *kss1Δ* single mutant does not (Fig. 1) (see also ref. 5). The potent MAPK-independent regulatory function of Kss1 has not been described previously for any MAPK, yet may be a general mechanism in MAPK signalling pathways that act as developmental switches. The molecular basis of Kss1 repression may involve its interactions with the proteins Dig1 and Dig2,

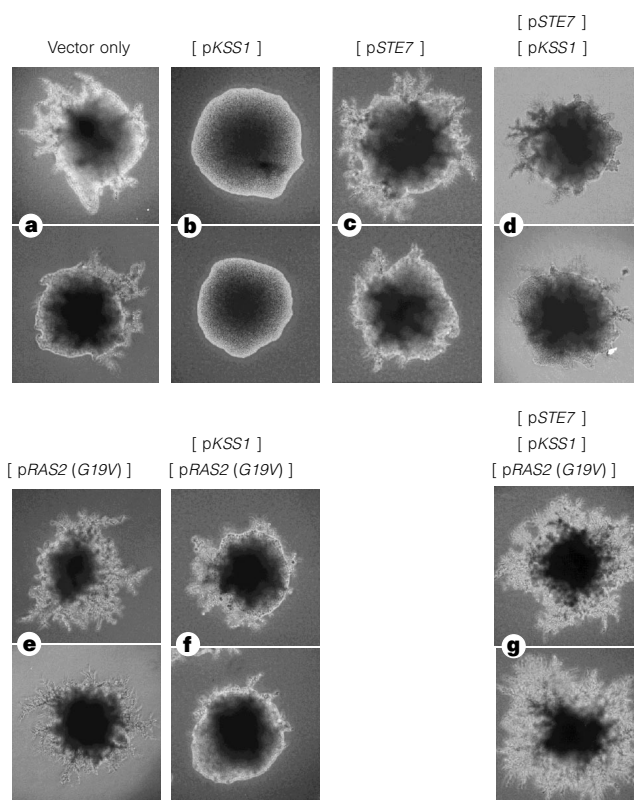


Figure 4 Ste7 and Kss1 potentiate, but are not required for, diploid pseudohyphal development in response to nutrient deprivation or activated Ras. Strain JCY119 (*MATa/MATα, ste7Δ/ste7Δ kss1Δ/kss1Δ*) was transformed with low-copy (CEN) plasmids, as follows: **a**, YCpU, YCpT and YCpL; **b**, YCpU, YCpT and YCpL-KSS1; **c**, YCpU, YCpT-STE7 and YCpL; **d**, YCpU, YCpT-STE7 and YCpL-KSS1; **e**, YCpU-RAS2^{G19V}, YCpT and YCpL; **f**, YCpU-RAS2^{G19V}, YCpT and YCpL-KSS1; and **g**, YCpU-RAS2^{G19V}, YCpT-STE7 and YCpL-KSS1. The resulting transformants were assayed for pseudohyphal filament formation on low-nitrogen plates as described¹⁹. Two representative colonies, photographed after 5 days of growth at 30°C, are shown for each strain.

which associate with Ste12, Kss1 and Fus3 (refs 16, 17). Finally, our observations suggest that, in haploids, invasive growth is the default developmental pathway and that the primary role of the pheromone-responsive MAPK cascade is to suppress invasive growth to allow haploids the developmental option of mating. Indeed, at least two pheromone-responsive genes, Fus3 (ref. 7) and Dig2 (ref. 16), are significant negative regulators of invasive growth^{5,16,17} and may act to reinforce further this developmental switch.

Methods

Yeast strain construction. Standard growth media and procedures for genetic manipulation of yeast have been described in detail¹⁸. All strains were constructed in the Σ 1278b background. JCY100 (*MATa leu2Δ::hisG his3Δ::hisG trp1Δ::hisG ura3-52*) and derivatives JCY110 (*kss1Δ::hisG*), JCY120 (*fus3Δ::TRP1*) and JCY130 (*kss1Δ::hisG fus3Δ::TRP1*), are identical to the corresponding strains described in ref. 13 and were a gift from H. Madhani and G. R. Fink. JCY117 (*ste7Δ::HIS3 kss1Δ::hisG*) was constructed from JCY110 by replacement of the endogenous STE7 locus with a PCR product containing HIS3 as described¹⁹. JCY127 (*ste7Δ::HIS3 fus3Δ::TRP1*) was constructed from strain JCY120 using the same procedure. JCY106 (*ste7Δ::URA3*) was constructed from JCY100 by replacement¹⁹ of the endogenous STE7 locus with a PCR product that contains URA3. JCY107 (*ste7Δ::ura3*) is a spontaneous Ura⁻ and 5-fluoro-orotic acid-resistant derivative of JCY106. JCY136 (*ste7Δ::URA3 kss1Δ::hisG fus3Δ::TRP1*) and JCY137 (*ste7Δ::ura3 kss1Δ::hisG fus3Δ::TRP1*) were constructed from strain

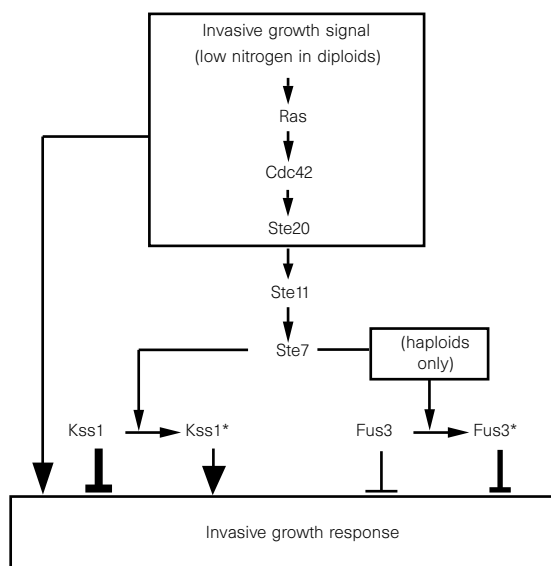


Figure 5 Model for regulation of haploid invasive growth and diploid pseudohyphal development by the Ste11-Ste7-Kss1 (and Fus3) MAPK cascade. Unactivated Kss1 functions as an inhibitor of invasive or filamentous growth in both haploids and diploids. Phosphorylation (*) by Ste7 switches Kss1 from an inhibitor to an activator. An efficacious MAPK-independent signal is also supplied through a Ras-dependent pathway that must branch upstream of Ste11 (see ref. 12). In haploids, Fus3 acts as a Ste7-dependent negative regulator of invasive growth⁵. Comparison of the *ste7Δ kss1Δ fus3Δ* triple mutant strain (JCY137) to its otherwise isogenic *FUS3** counterpart (JCY117), for either FRET-dependent transcription (Fig. 1e) or by following the time course of agar penetration (data not shown), showed that Fus3 also functions as a repressor, albeit weakly, even in the absence of activation by Ste7.

JCY130 using the same procedure. YLB537 (*ste11Δ::URA3 ste7Δ::ura3 kss1Δ::hisG fus3Δ::TRP1*) was constructed from JCY137 by deletion of *STE11* using a procedure described previously²⁰. YLB637 (*ste7Δ::ura3 kss1Δ::hisG fus3::TRP1 ste12Δ::LEU2*) was constructed from JCY137 by deletion of *STE12* using a method described before¹⁶. JCY119 (*MATa/MATα ste7Δ::HIS3 ste7Δ::HIS3 kss1Δ::hisG/kss1Δ::hisG*) was constructed by transforming strain JCY117 with two plasmids expressing the *HO* gene²¹ and *STE7*, respectively, and then screening for non-mating (presumably diploid) clones. After segregation of the plasmids under non-selective conditions, ploidy was confirmed by direct visualization of tetrads induced on sporulation medium¹⁸.

Plasmid construction. YCpT, YCpU and YCpL are YCplac22, YCplac33, and YCplac111, respectively²². YCpT-*STE7* (ref. 20) and YCpU-RAS2^{G19V} (refs 10, 12) have been described. To construct plasmid YCpL-KSS1, an *EcoRI*-*SphI* fragment containing *KSS1* and its promoter was excised from plasmid YEp-KSS1 (ref. 8) and inserted into the corresponding sites in YCplac111. The same scheme was used to derive YCpU-KSS1(K42R Q45P) and YCpU-KSS1(T183A Y185F) from their YEp-based counterparts⁸. To construct plasmid YEpU-FTyZ, the *CYC1* upstream activation sequence (UAS) was excised from plasmid pLG669-Z²³ by digestion with *XhoI* and *NheI*, and replaced with a double-stranded oligonucleotide containing the FRET from the transposon, Ty1 (ref. 13), generated by annealing Ty1FREw (5'-CTA GCA TTC TTC TGT TTT GGA AGC TGA AAC GC-3') to Ty1FREc (5'-TCG AGC GTT TCA GCT TCC AAA ACA GAA GAA TG-3').

Assays for mating, invasive growth, filament formation, and β -galactosidase activity. Patch mating assays were performed as described²⁴. Assays for haploid invasive growth⁵ and diploid pseudohyphal development^{6,10} were performed as described. For determination of β -galactosidase activity, cells were grown to mid-exponential phase in synthetic minimal medium¹⁸ without uracil containing 2% glucose and protein extracts were prepared²⁰; either 5 μ g (for high expressers) or 200 μ g (for low expressers) of total protein was assayed as described previously²⁰.

Received 11 August; accepted 15 September 1997.

- Bardwell, L., Cook, J. G., Inouye, C. J. & Thorner, J. Signal propagation and regulation in the mating pheromone response pathway of the yeast *Saccharomyces cerevisiae*. *Dev. Biol.* **166**, 363–379 (1994).
- Cobb, M. H. & Goldsmith, E. J. How MAP kinases are regulated. *J. Biol. Chem.* **270**, 14843–14846 (1995).
- Denhardt, D. T. Signal-transducing protein phosphorylation cascades mediated by Ras/Rho proteins in the mammalian cell—the potential for multiplex signaling. *Biochem. J.* **318**, 729–747 (1996).
- Zhou, Z., Gartner, A., Cade, R., Ammerer, G. & Errede, B. Pheromone-induced signal transduction in *Saccharomyces cerevisiae* requires the sequential function of three protein kinases. *Mol. Cell. Biol.* **13**, 2069–2080 (1993).
- Roberts, R. L. & Fink, G. R. Elements of a single MAP kinase cascade in *Saccharomyces cerevisiae* mediate two developmental programs in the same cell type: mating and invasive growth. *Genes Dev.* **8**, 2974–2985 (1994).
- Liu, H., Styles, C. A. & Fink, G. R. Elements of the yeast pheromone response pathway required for filamentous growth of diploids. *Science* **262**, 1741–1744 (1993).
- Elion, E. A., Brill, J. A. & Fink, G. R. *FUS3* represses *CLN1* and *CLN2* and in concert with *KSS1* promotes signal transduction. *Proc. Natl Acad. Sci. USA* **88**, 9392–9396 (1991).
- Ma, D., Cook, J. G. & Thorner, J. Phosphorylation and localization of Kss1, a MAP kinase of the *Saccharomyces cerevisiae* pheromone response pathway. *Mol. Biol. Cell* **6**, 889–909 (1995).
- Hunter, T. & Plowman, G. D. The protein kinases of budding yeast: six score and more. *Trends Biochem. Sci.* **22**, 18–22 (1997).
- Gimeno, C. J., Ljungdahl, P. O., Styles, C. A. & Fink, G. R. Unipolar cell divisions in the yeast *S. cerevisiae* lead to filamentous growth: regulation by starvation and RAS. *Cell* **68**, 1077–1090 (1992).
- Gavrias, V., Andrianopoulos, A., Gimeno, C. J. & Timberlake, W. E. *Saccharomyces cerevisiae* TEC1 is required for pseudohyphal growth. *Mol. Microbiol.* **19**, 1255–1263 (1996).
- Mösch, H. U., Roberts, R. L. & Fink, G. R. Ras2 signals via the Cdc42/Ste20/mitogen-activated protein kinase module to induce filamentous growth in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **93**, 5352–5356 (1996).
- Madhani, H. D. & Fink, G. R. Combinatorial control required for the specificity of yeast MAPK signaling. *Science* **275**, 1314–1317 (1997).
- Neiman, A. M. & Herskowitz, I. Reconstitution of a yeast protein kinase cascade *in vitro*: activation of yeast MEK homologue STE7 by STE11. *Proc. Natl Acad. Sci. USA* **91**, 3398–3402 (1994).
- Courchesne, W. E., Kunisawa, R. & Thorner, J. A putative protein kinase overcomes pheromone-induced arrest of cell cycling in *S. cerevisiae*. *Cell* **58**, 1107–1119 (1989).
- Cook, J. G., Bardwell, L., Kron, S. J. & Thorner, J. Two novel targets of the MAP kinase Kss1 are negative regulators of invasive growth in the yeast *Saccharomyces cerevisiae*. *Genes Dev.* **10**, 2831–2848 (1996).
- Tedford, K., Kim, S., Sa, D., Stevens, K. & Tyers, M. Regulation of the mating pheromone and invasive growth responses in yeast by two MAP kinase substrates. *Curr. Biol.* **7**, 228–238 (1997).
- Sherman, F., Fink, G. R. & Hicks, J. B. *Methods in Yeast Genetics* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1986).
- Lorenz, M. C. Gene disruption with PCR products in *Saccharomyces cerevisiae*. *Gene* **158**, 113–117 (1995).
- Bardwell, L., Cook, J. G., Chang, E. C., Cairns, B. R. & Thorner, J. Signaling in the yeast pheromone response pathway: specific and high-affinity interaction of the mitogen-activated protein (MAP) kinases Kss1 and Fus3 with the upstream MAP kinase kinase Ste7. *Mol. Cell. Biol.* **16**, 3637–3650 (1996).
- Herskowitz, I. & Jensen, R. E. Putting the HO gene to work: practical uses for mating-type switching. *Meth. Enzymol.* **194**, 132–146 (1991).
- Gietz, R. D. & Sugino, A. New yeast-*Escherichia coli* shuttle vectors constructed with *in vitro* mutagenized yeast genes lacking six-base pair restriction sites. *Gene* **74**, 527–534 (1988).
- Guarante, L. & Ptashne, M. Fusion of *Escherichia coli lacZ* to the cytochrome c gene of *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **78**, 2199–2203 (1981).
- Sprague, G. F. Jr Assay of yeast mating reaction. *Meth. Enzymol.* **194**, 77–93 (1991).

Acknowledgements. We thank D. Voora for technical assistance with the data shown in Fig. 1e, and G. R. Fink for reagents. This work was supported by an NIH-NIGMS predoctoral traineeship and an NIH-NCI postdoctoral traineeship (to J.G.C.), an NIH-NCI postdoctoral traineeship and a senior postdoctoral fellowship from the Leukemia Society of America (to L.B.), an NIH-NIGMS research grant (to J.T.) and resources provided by the Berkeley campus Cancer Research Laboratory.

Correspondence and requests for materials should be addressed to J.T. (e-mail: jthorner@mendel.berkeley.edu).

Switching of the coupling of the β_2 -adrenergic receptor to different G proteins by protein kinase A

Yehia Daaka, Louis M. Luttrell & Robert J. Lefkowitz

Howard Hughes Medical Institute, Departments of Medicine and Biochemistry, Box 3821, Duke University Medical Center, Durham, North Carolina 27710, USA

Many of the G-protein-coupled receptors for hormones that bind to the cell surface can signal to the interior of the cell through several different classes of G protein^{1–4}. For example, although most of the actions of the prototype β_2 -adrenergic receptor are mediated through G_s proteins and the cyclic-AMP-dependent protein kinase (PKA) system^{5,6}, β_2 -adrenergic receptors can also couple to G_i proteins^{1,2}. Here we investigate the mechanism that controls the specificity of this coupling. We show that in HEK293 cells, stimulation of mitogen-activated protein (MAP) kinase by the β_2 -adrenergic receptor is mediated by the $\beta\gamma$ subunits of pertussis-toxin-sensitive G proteins through a pathway involving

the non-receptor tyrosine kinase c-Src and the G protein Ras. Activation of this pathway by the β_2 -adrenergic receptor requires that the receptor be phosphorylated by PKA because it is blocked by H-89, an inhibitor of PKA. Additionally, a mutant of the receptor, which lacks the sites normally phosphorylated by PKA, can activate adenylyl cyclase⁵, the enzyme that generates cAMP, but not MAP kinase. Our results demonstrate that a mechanism previously shown to mediate uncoupling of the β_2 -adrenergic receptor from G_s and thus heterologous desensitization⁷ (PKA-mediated receptor phosphorylation), also serves to 'switch' coupling of this receptor from G_s to G_i and initiate a new set of signalling events.

Stimulation by isoprenaline (isoproterenol) of endogenous β_2 -adrenergic receptors (β -ARs) in HEK293 cells results in a sixfold increase in the phosphorylation state of MAP kinase (Fig. 1a, b). This increase is saturable (half-maximal effective concentration (EC_{50}) of 10^{-8} M), in agreement with results reported for Cos-7 cells⁸. The isoprenaline-stimulated MAP kinase activation is mediated by β -ARs, because addition of propranolol, a β -AR antagonist, abrogates the ligand-mediated increase in active MAP kinase (Fig. 1c), whereas addition of prazosin, an α_1 -AR antagonist, has no effect. Surprisingly, pretreatment of HEK293 cells with pertussis toxin (PTX) abolishes the β -AR-mediated increase in MAP kinase phosphorylation (Fig. 2a). Transient overexpression of a G $\beta\gamma$ sequesterant peptide derived from the carboxy terminus of β -adrenergic receptor kinase 1 (β -ARKct⁹) also markedly inhibited MAP kinase phosphorylation (Fig. 2a). These results indicate that β -AR-mediated MAP kinase activation is mediated by G $\beta\gamma$ subunits derived from PTX-sensitive G proteins.

In many cells, G_i -coupled receptors mediate G $\beta\gamma$ -subunit-dependent MAP kinase activation through a pathway involving the Src family of tyrosine kinases^{10–12} and the low-molecular-mass G protein Ras^{13,14}. As shown in Fig. 2b, transient overexpression of p50^{csk}, an inhibitor of Src^{10,11}, or Sos-Pro, an inhibitor of the Ras guanine-nucleotide-exchange factor mSos1 (ref. 15), causes a 50–60% decrease in β -AR-mediated MAP kinase phosphorylation, suggesting that the β -AR-mediated pathway for activation of MAP kinase is identical to the previously described G_i -coupled-receptor-mediated pathway^{8,11,15}.

To determine whether β -AR stimulation directly results in G_i activation, plasma membranes isolated from HEK293 cells were stimulated with isoprenaline in the presence of a photoactivatable [γ -³²P]GTP analogue. After stimulation, G_i proteins were immunoprecipitated and the GTP incorporated was detected by autoradiography. Figure 2c shows the specific GTP loading of G_i following isoprenaline stimulation. These results support a role for G_i coupling to the β -AR in signal transduction from the receptor to MAP kinase.

Most of the metabolic effects of β -AR stimulation result from G_s -mediated activation of adenylyl cyclases and activation of PKA^{5,6}. In human endothelial cells, downregulation of $G_s\alpha$ abolishes β -AR-mediated activation of MAP kinase by isoprenaline¹⁶, suggesting that G_s activation is required for MAP kinase activation by this receptor. To determine whether PKA activation is a prerequisite for β -AR coupling to G_i , we tested the effects of PKA inhibition on MAP kinase activation in HEK293 cells. As shown in Fig. 3, pretreatment of cells with H-89, a specific, cell-permeable inhibitor of PKA¹⁷, inhibited isoprenaline-mediated activation of MAP kinase, with no effect on activation mediated by the G_i -coupled receptor for lysophosphatidic acid (LPA). Moreover, inhibition of the PKA activity attenuated the G_i incorporation of GTP following isoprenaline treatment (Fig. 2c) but had no effect on the LPA-mediated loading of G_i (data not shown).

PKA-mediated phosphorylation of the β -AR is known to impair receptor- G_s coupling^{5,6}. A β -AR mutant lacking the putative PKA phosphorylation sites (β -AR_{mut}) stimulates adenylyl cyclase as effectively as wild-type β -AR, but does not undergo PKA-mediated

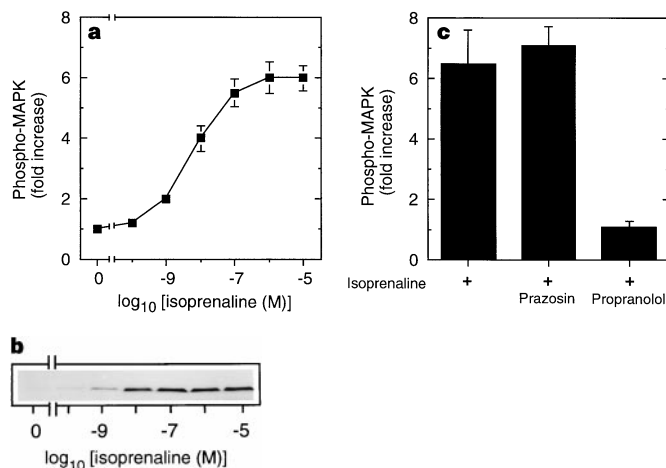


Figure 1 β -AR-mediated activation of MAP kinase. **a**, Dose-dependent activation of MAPK by isoprenaline. Serum-starved HEK293 cells were incubated with increasing concentrations of isoprenaline for 5 min and then the amount of agonist-induced phospho-MAPK was determined. **b**, Representative immunoblot showing MAPK phosphorylated state. **c**, Isoprenaline-induced phosphorylation of MAPK is mediated by endogenous β -AR. Cells were preincubated with the α_1 -AR antagonist prazosin (1 μ M) or the β -AR antagonist propranolol (10 μ M) for 10 min before stimulation with isoprenaline (1 μ M). Active MAPK was assayed by using anti-phospho-MAPK antibodies; values represent means $\pm$ s.e. from 3 independent experiments.

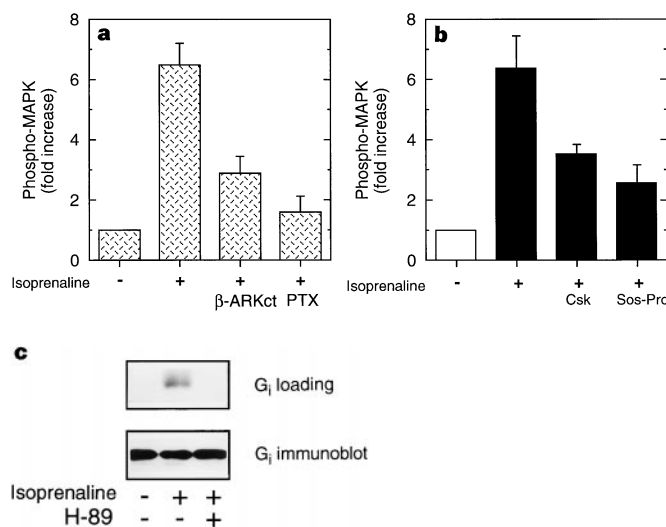


Figure 2 Isoprenaline activation of MAPK is dependent on PTX-sensitive G proteins. **a**, Accumulation of phospho-MAPK following agonist stimulation is inhibited by β -ARKct and PTX. Cells transiently expressing β -ARKct were starved overnight in the presence or absence of PTX (100 ng ml⁻¹) before determining agonist-mediated phosphorylation of MAPK. **b**, Effects of c-Src- and Sos-inhibitory proteins on phospho-MAPK by isoprenaline. Cells were transiently transfected with p50^{csk} or Sos-Pro, stimulated with isoprenaline, and active MAPK levels determined. Data represent the isoprenaline-induced fold increase of MAPK phosphorylation relative to unstimulated cells. **c**, PKA mediates the β -AR-induced GTP loading of G_i . Membranes were photolabelled after preincubation with [γ -³²P]GTP azidoanilide in the absence (–) or presence (+) of isoprenaline or H-89. $G_i\alpha$ proteins were immunoprecipitated, fractionated on SDS-PAGE and autoradiographed. Lower panel, the same filter used for the G_i -GTP crosslinking was used as a blot to test for the presence of immunoprecipitated $G_i\alpha$ proteins.

desensitization⁵. To determine whether PKA-mediated phosphorylation of the β -AR affects the ability of the receptor to signal through G_i , we tested whether this β -AR_{mut} could stimulate MAP kinase phosphorylation. Isoprenaline-stimulated MAP kinase phosphorylation was determined in HEK293 cells transiently overexpressing an epitope-tagged Erk1 (p44^{HA-mapk}) plus either wild-type β -AR or β -AR_{mut}. Both wild-type and mutant receptors were overexpressed to levels of 300–400 fmol per mg whole-cell protein. As shown in Fig. 4, overexpression of wild-type receptor results in 40–50% enhancement of p44^{HA-mapk} phosphorylation above that mediated by endogenous β -ARs in control cells. In cells expressing the β -AR_{mut}, isoprenaline-stimulated phosphorylated p44^{HA-mapk} was inhibited compared to controls. Thus, the PKA-phosphorylation-site mutant β -AR not only failed to stimulate G_i -mediated MAP kinase activation, but behaved as a dominant suppressor of the isoprenaline-mediated MAP kinase signalling pathway triggered by endogenous β -AR.

Collectively, our data suggest a model for β -AR-mediated MAP kinase activation in which PKA-mediated phosphorylation of the receptor enables it to couple to PTX-sensitive G proteins. After phosphorylation by PKA, a peptide that is derived from the third intracellular loop of β_2 -AR and which contains a putative PKA-phosphorylation site is less able to stimulate G_s , but can stimulate G_i more *in vitro*¹⁸. The following model is shown in Fig. 5: upon agonist binding, the receptor-catalysed activation of G_s stimulates adenylyl cyclase activity; the resulting rise in intracellular cAMP activates PKA and the agonist-occupied receptor is phosphorylated. After phosphorylation, the receptor switches its coupling specificity to G_i . GTP-bound $G_i\alpha$ dissociates from the heterodimeric $G\beta\gamma$, and free $G\beta\gamma$ subunits mediate activation of the MAP kinase signalling pathway in the same way as G_i -coupled receptors.

Depending on cell type, an increase in intracellular cAMP may result in cellular proliferation^{19,20}, differentiation^{21,22} or growth arrest²³. cAMP-dependent inhibition of Raf1 antagonizes MAP kinase activation in fibroblasts^{24,25} at a step in the signalling cascade after Ras activation. In contrast, an increase in intracellular cAMP in Cos-7 cells by a route not involving a receptor, including expression of a constitutively active mutant $G_s\alpha$, or exposure to 8-bromo-cAMP or forskolin, results in MAP kinase activation^{8,26}. In HEK293 cells, treatment with forskolin or isoprenaline increases cAMP by 45 ± 2.2 and 17 ± 1.5 times, respectively. Forskolin treatment induces an (8.2 ± 0.6) -fold increase in MAP kinase phosphorylation, suggesting that cAMP may mediate MAP kinase activation by a receptor-independent pathway. However, forskolin-stimulated MAP kinase phosphorylation in HEK293 cells is insensitive to both PTX and β -ARKct expression, indicating that forskolin works by a different mechanism from that used by the β -AR.

By coupling to different classes of G protein, receptors can

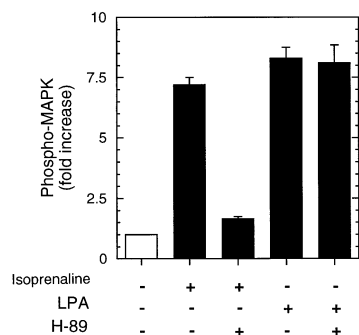


Figure 3 Effect of the cAMP-dependent protein kinase A inhibitor H-89 on MAP kinase activation. HEK293 cells were incubated overnight in serum-free media before 30 min pretreatment with $10 \mu\text{M}$ H-89. Isoprenaline ($1 \mu\text{M}$) or LPA ($10 \mu\text{M}$) were added to samples for 5 min in the presence of H-89 and active MAPK levels determined. Values shown represent means $\pm$ s.e. from 3 experiments and are expressed as fold increase over unstimulated cells.

simultaneously activate several distinct signalling cascades. Examples of such dual coupling of β -AR receptors to G_s and G_i pathways have been reported^{1,2}. PTX treatment of cardiac myocytes enhances β_2 -AR-stimulated increases in Ca^{2+} current, Ca_i transients, and contraction amplitude². Our results indicate that the coupling of β -AR to G_i must be a regulated process, because the β -AR must be phosphorylated by PKA before PTX-sensitive MAP kinase activation can occur.

Together, our results indicate that PKA phosphorylation of the β -AR plays a key regulatory role in β -AR signalling. By decreasing receptor-coupling efficiency to G_s , PKA phosphorylation brings about heterologous receptor desensitization, reducing cAMP production in response to further stimulation. Simultaneously, the

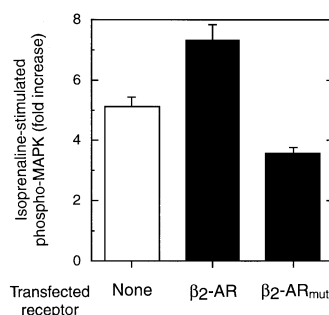


Figure 4 PKA phosphorylation of the β -AR is required for isoprenaline-induced phosphorylation of MAPK. Cells were transiently transfected with haemagglutinin-tagged Erk1 (p44^{HA-mapk}) either alone or with wild-type β_2 -AR, or with β_2 -AR lacking phosphorylation sites for PKA (β_2 -AR_{mut}). Cells were starved overnight before assaying for isoprenaline-induced MAPK phosphorylation. Values shown represent means $\pm$ s.e. of phosphorylated p44^{HA-mapk} and are expressed as fold increase over unstimulated cells. The expression of receptor was as follows: endogenous β_2 -AR, 30 fmol mg^{-1} protein; overexpressed β_2 -AR or β_2 -AR_{mut}, 300 – 400 fmol mg^{-1} protein. Values for the isoprenaline-induced increase of intracellular cAMP were as follows: endogenous β_2 -AR, (15.7 ± 0.6) -fold; overexpressed β_2 -AR, (28.5 ± 1.2) -fold; overexpressed β_2 -AR_{mut}, (29.6 ± 1.3) -fold relative to unstimulated cells.

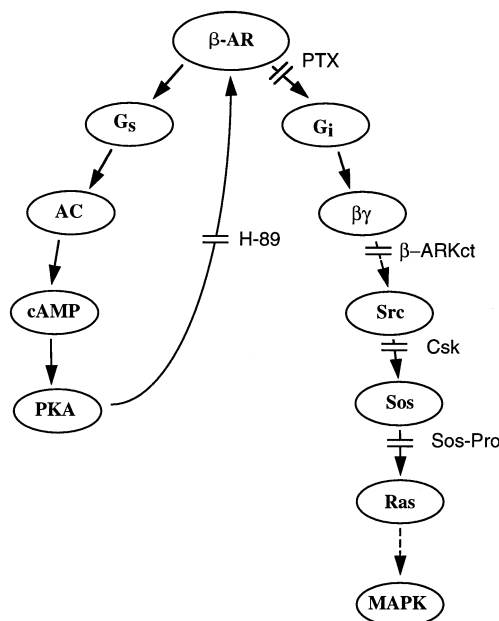


Figure 5 Model for β_2 -AR-mediated G-protein switching to activate MAP kinase. Phosphorylation of MAP kinase by the β -AR proceeds through receptor coupling to G_s and activation of PKA. Activated PKA phosphorylates the β -AR leading to receptor coupling to, and activation of, G_i . $G\beta\gamma$ released from the G_i -coupled β -AR activates MAP kinase in a Src- and Ras-dependent pathway.

enhanced coupling efficiency to G_i results in direct feedback inhibition of the G_s -mediated adenylyl cyclase signal that initiated the process. Moreover, coupling of the β -AR to G_i initiates a second wave of G_i -mediated signalling, including the rapid activation of MAP kinases, which coincides with termination of the signal that initiated the original response. Thus, PKA-mediated phosphorylation of the β -AR serves as a switching mechanism to regulate the G-protein-coupling specificity of the receptor. □

Methods

Cell culture and transfection. The cDNAs encoding wild-type β_2 -AR, β_2 -AR lacking putative phosphorylation sites for cAMP-dependent protein kinase A (β_2 -AR^{mut}; point mutations of serine residues at 261, 262, 345 and 346 to alanine), the β -ARK C-terminal peptide-encoding minigene (β -ARKct) and the C-terminal fragment of mSos1 (Sos-Pro) were constructed in our laboratory^{5,9,15}. The cDNAs encoding p50^{sk} and haemagglutinin-tagged p44^{mapk} (p44^{HA-mapk}) were provided by H. Hanafusa and J. Pouyssegur, respectively. Cells were collected and re-established in fresh medium 24 h before each experiment. Transient transfections were done using calcium phosphate co-precipitation²⁷. Cells were starved overnight in medium containing 10 mM HEPES (pH 7.4) and 0.1% (v/v) BSA before agonist stimulation. All assays were done 48 h after transfection. Transient expression of plasmids transfected with p50^{sk}, mutant mSos1 or β -ARKct was verified by immunoblotting of whole-cell lysates using commercially available antibodies.

Phospho-MAPK assay. HEK293 cells were treated with isoprenaline (1 μ M) or lysophosphatidic acid (10 μ M) for 5 min at 37 °C. Cell monolayers were lysed in Laemmli sample buffer and lysates resolved by SDS-PAGE. Activated MAP kinase on nitrocellulose was detected using anti-phospho-MAP-kinase antibodies (New England BioLabs), visualized by enzyme-linked chemiluminescence (Amersham) and quantified by scanning laser densitometry.

G_i loading. Membranes²⁸ from HEK293 cells were aliquoted and stored at -70 °C in TE (20 mM Tris, pH 7.4, 1 mM EDTA) buffer. Membrane proteins (40–50 μ g) were preincubated in TE buffer containing 50 mM CH₃CO₂Na, 0.2 mM EGTA, 1.0 mM benzimidazole, 2 mM MgCl₂ and 50 μ M GDP in order to load G α subunits. Isoprenaline and [γ -³²P]GTP-azidoanilide (Research Products) were added and samples incubated for 10 min at ambient temperature. The reaction was terminated by centrifugation (14,000g for 10 min) and membrane pellets were resuspended in TE containing 1 mM dithiothreitol and irradiated for 10 min with an ultraviolet lamp (λ = 254 nm). Membranes were collected by centrifugation and solubilized in 40 μ l of 2% SDS followed by the addition of 800 μ l of RIPA buffer (150 mM NaCl, 50 mM Tris, pH 8.0, 5 mM EDTA, 1 mM dithiothreitol, 1% Nonidet P-40, 0.5% deoxycholate, 10 mM NaF, 1 mM Na₃VO₄, 0.2 mM PMSF and 10 μ g aprotinin per ml). Samples were incubated with 15 μ g anti-G α antiserum and 50 μ l protein A/G-Sepharose beads and rotated overnight at 4 °C. Beads were washed three times with ice-cold RIPA buffer, resuspended in Laemmli buffer and separated by SDS-PAGE. Proteins were transferred onto nitrocellulose filter and analysed for [γ -³²P]GTP incorporation by autoradiography. Filters were also analysed for immunoprecipitation efficiency by blotting for the presence of G α proteins.

Received 28 April; accepted 6 August 1997.

- Abramson, S. *et al.* Interaction of β -adrenergic receptors with the inhibitory guanine nucleotide-binding protein of adenylyl cyclase in membranes prepared from cyc^c S49 lymphoma cells. *Biochem. Pharmacol.* **37**, 4289–4297 (1988).
- Xiao, R., Ji, X. & Lakatta, E. G. Functional coupling of the β_2 -adrenoceptor to a pertussis toxin-sensitive G protein in cardiac myocytes. *Mol. Pharmacol.* **47**, 322–329 (1995).
- Laugwitz, K.-L. *et al.* The human thyrotropin receptor: A heptahelical receptor capable of stimulating members of all four G protein families. *Proc. Natl Acad. Sci. USA* **93**, 116–120 (1996).
- Herrlich, A. *et al.* Involvement of G α_i and G α_o proteins in dual coupling of the luteinizing hormone receptor to adenylyl cyclase and phospholipase C. *J. Biol. Chem.* **271**, 16764–16772 (1996).
- Hausdorff, W. P. *et al.* Phosphorylation sites on two domains of the β_2 -adrenergic receptor are involved in distinct pathways of receptor desensitization. *J. Biol. Chem.* **264**, 12657–12665 (1989).
- Strulovici, B., Cerione, R. A., Kilpatrick, B. F., Caron, M. G. & Lefkowitz, R. J. Direct demonstration of impaired functionality of a purified desensitized β -adrenergic receptor in a reconstituted system. *Science* **225**, 837–840 (1984).
- Freedman, N. J. & Lefkowitz, R. J. Desensitization of G protein-coupled receptors. *Rec. Progr. Horm. Res.* **51**, 319–353 (1996).
- Crespo, P., Cachero, T. G., Xu, N. & Gutkind, J. S. Dual effect of β -adrenergic receptors on mitogen-activated protein kinase: Evidence for a β - γ -dependent activation and G α -cAMP-mediated inhibition. *J. Biol. Chem.* **270**, 25259–25265 (1995).
- Koch, W. J., Ingles, J., Stone, W. C. & Lefkowitz, R. J. The binding site for β subunits of heterotrimeric G proteins on the β -adrenergic receptor kinase. *J. Biol. Chem.* **268**, 8256–8260 (1993).
- Luttrell, L. M. *et al.* Role of c-Src tyrosine kinase in G protein-coupled receptor- and G $\beta\gamma$ subunit-mediated activation of mitogen-activated protein kinases. *J. Biol. Chem.* **271**, 19443–19450 (1996).

- Luttrell, L. M., Della Rocca, G. J., van Biesen, T., Luttrell, D. K. & Lefkowitz, R. J. G $\beta\gamma$ subunits mediate Src-dependent phosphorylation of the epidermal growth factor receptor: A scaffold for G protein-coupled receptor-mediated Ras activation. *J. Biol. Chem.* **272**, 4637–4644 (1997).
- Chen, Y., Pouyssegur, J., Courtneidge, S. A. & van Obberghen-Schilling, E. Activation of Src family kinase activity by G protein-coupled thrombin receptor in growth-responsive fibroblasts. *J. Biol. Chem.* **269**, 27372–27377 (1994).
- Boguski, M. S. & McCormick, F. Proteins regulating Ras and its relatives. *Nature* **336**, 643–653 (1993).
- Hordijk, P. L., Verlaan, L., van Corven, E. J. & Moolenaar, W. H. Protein tyrosine phosphorylation induced by lysophosphatidic acid in Rat-1 fibroblasts: Evidence that phosphorylation of MAP kinase is mediated by the Gi-p21ras pathway. *J. Biol. Chem.* **269**, 645–651 (1994).
- van Biesen, T. *et al.* Receptor-tyrosine kinase- and G $\beta\gamma$ -mediated MAP kinase activation by a common signalling pathway. *Nature* **376**, 781–784 (1995).
- Sextl, V. *et al.* Stimulation of the mitogen-activated protein kinase via the A_{2A}-adenosine receptor in primary human endothelial cells. *J. Biol. Chem.* **272**, 5792–5799 (1997).
- Chijiwa, T. *et al.* Inhibition of forskolin-induced neurite outgrowth and protein phosphorylation by a newly synthesized selective inhibitor of cyclic-AMP-dependent protein kinase, N-[2-(p-bromocinnamylamino)ethyl]-5-isoquinolinesulfonamide (H-89), of PC12D pheochromocytoma cells. *J. Biol. Chem.* **265**, 5267–5272 (1990).
- Okamoto, T. *et al.* Identification of a Gs activator region of the β_2 -adrenergic receptor that is autoregulated via protein kinase A-dependent phosphorylation. *Cell* **67**, 723–730 (1991).
- Chambard, J. C., Paris, L., L'Allemand, G. & Pouyssegur, J. Two growth factor signalling pathways in fibroblasts distinguished by pertussis toxin. *Nature* **326**, 800–803 (1987).
- van Corven, E. J., Groenink, A., Jalik, K., Eichholtz, T. & Moolenaar, W. H. Lysophosphatidate-induced cell proliferation: identification and dissection of signaling pathways mediated by G proteins. *Cell* **59**, 45–54 (1989).
- Cowley, S., Paterson, H., Kemp, P. & Marshall, C. J. Activation of MAP kinase kinase is necessary and sufficient for PC12 differentiation and transformation of NIH3T3 cells. *Cell* **77**, 841–852 (1994).
- Yarwood, S. J., Kilgour, E. & Anderson, N. G. Cyclic AMP stimulates the phosphorylation and activation of p42 and p44 mitogen-activated protein kinases in 3T3-F442A preadipocytes. *Biochem. Biophys. Res. Comm.* **224**, 734–739 (1996).
- McKenzie, F. R. & Pouyssegur, J. cAMP-mediated growth inhibition in fibroblasts is not mediated via mitogen-activated protein (MAP) kinase (Erk) inhibition: cAMP-dependent protein kinase induces a temporal shift in growth factor-stimulated MAP kinases. *J. Biol. Chem.* **271**, 13476–13483 (1996).
- Cook, S. J. & McCormick, F. Inhibition by cAMP of Ras-dependent activation of Raf. *Science* **262**, 1069–1072 (1993).
- Hordijk, P. L., Verlaan, L., Jalnik, K., van Corven, E. J. & Moolenaar, W. H. cAMP abrogates the p21^{ras}-mitogen-activated protein kinase pathway in fibroblasts. *J. Biol. Chem.* **269**, 3534–3538 (1994).
- Faure, M., Voinyo-Yasenetskaya, T. A. & Bourne, H. R. cAMP and $\beta\gamma$ subunits of heterotrimeric G proteins stimulates the mitogen-activated protein kinase pathway in Cos-7 cells. *J. Biol. Chem.* **269**, 7851–7854 (1994).
- Daaka, Y. *et al.* Receptor and G $\beta\gamma$ isoform-specific interactions with G protein-coupled receptor kinases. *Proc. Natl Acad. Sci. USA* **94**, 2180–2185 (1997).
- Krueger, K. M., Daaka, Y., Pitcher, J. A. & Lefkowitz, R. J. The role of sequestration in G protein-coupled receptor resensitization: Regulation of β_2 -adrenergic receptor dephosphorylation by vesicular acidification. *J. Biol. Chem.* **272**, 5–8 (1997).

Acknowledgements. We thank M. Holben and D. Addison for secretarial assistance.

Correspondence and requests for materials should be addressed to R.J.L. (e-mail: lefko001@mc.duke.edu).

Tyrosine phosphorylation of the EGF receptor by the kinase Jak2 is induced by growth hormone

Toshimasa Yamauchi*†, Kohjiro Ueki*†, Kazuyuki Tobe*, Hiroyuki Tamemoto*, Nobuo Sekine‡, Mitsufumi Wada*, Masaru Honjo§, Michio Takahashi||, Tokiharu Takahashi*, Hisamaru Hirai*, Toshio Tushima¶, Yasuo Akanuma#, Toshiro Fujita‡, Issei Komuro*, Yoshio Yazaki* & Takashi Kadowaki*

* Third Department of Internal Medicine, Faculty of Medicine, University of Tokyo, Tokyo 113, Japan

‡ Fourth Department of Internal Medicine, Faculty of Medicine, University of Tokyo, Tokyo 112, Japan

§ Life Science Laboratories, Mitsui Toatsu Chemicals Inc., Mobara 297, Japan

|| Department of Veterinary Physiology, Veterinary Medical Science, University of Tokyo, Tokyo 113, Japan

¶ Second Department of Internal Medicine, Tokyo Women's Medical College, Tokyo 162, Japan

Institute for Diabetes Care and Research, Asahi Life Foundation, Tokyo 100, Japan

† These authors contributed equally to this work.

When growth hormone binds to its receptor, which belongs to the cytokine receptor superfamily¹, it activates the Janus kinase Jak2² which has tyrosine-kinase activity and initiates an activation of several key intracellular proteins (for example, mitogen-activated protein (MAP) kinases^{3–6}) that eventually execute the biological

actions induced by growth hormone, including the expression of particular genes. In contrast to receptors that themselves have tyrosine kinase activity, the signalling pathways leading to MAP kinase activation^{7,8} that are triggered by growth hormone are poorly understood, but appear to be mediated by the proteins Grb2 and Shc⁹. We now show that growth hormone stimulates tyrosine phosphorylation of the receptor for epidermal growth factor (EGFR) and its association with Grb2 and at the same time stimulates MAP kinase activity in liver, an important target tissue of growth hormone. Expression of EGFR and its mutants revealed that growth-hormone-induced activation of MAP kinase and expression of the transcription factor *c-fos* requires phosphorylation of tyrosines on EGFR, but not its own intrinsic tyrosine-kinase activity. Moreover, tyrosine at residue 1,068 of the EGFR is proposed to be one of the principal phosphorylation sites and Grb2-binding sites stimulated by growth hormone via Jak2. Our results indicate that the role of EGFR in signalling by growth hormone is to be phosphorylated by Jak2, thereby providing docking sites for Grb2 and activating MAP kinases and gene expression, independently of the intrinsic tyrosine kinase activity of EGFR. This may represent a novel cross-talk pathway between the cytokine receptor superfamily and growth factor receptor.

Insulin-receptor substrates IRS-1 and IRS-2 are tyrosine-phosphorylated in response to growth hormone (GH) in cultured cells such as primary rat adipocytes, 3T3-F442A fibroblasts, and CHO-GH receptor (GHR) cells¹⁰⁻¹². We detected GH-dependent tyrosyl phosphorylation of a protein with a relative molecular mass (M_r) of 180K (pp180), Jak2 and GHR (Fig. 1a, top, lane 3, and data not shown) in mouse liver, a major target tissue, *in vivo*: its maximal

tyrosine phosphorylation was roughly equivalent to that of IRS-1 (M_r 170K) observed following treatment with insulin (Fig. 1a, top, lane 2). IRS-1 deficiency¹³⁻¹⁵ had no apparent effect on GH-induced tyrosine phosphorylation of pp180 (Fig. 1a, top, lane 6). In addition, although IRS-1 deficiency induced tyrosine phosphorylation of IRS-2 (M_r ~180K) instead of IRS-1 in response to insulin (Fig. 1a, top, lane 5), immunodepletion of IRS-2 using specific antibody did not cause a significant decrease in tyrosine phosphorylation of pp180 in response to GH (data not shown). Thus the identity of the pp180, the principal protein tyrosine-phosphorylated in response to GH, is not IRS-1/IRS-2. GH stimulated association of pp180 with Grb2 but not with p85 phosphatidylinositol-3-OH kinase (PI(3)K) (data not shown). In addition, we were unable to extract pp180 without detergent, and found that it bound to wheat germ agglutinin-linked beads. We therefore tested whether EGFR could be pp180 by stripping and reprobing the sheet in Fig. 1a (top) with anti-EGFR antibody. We found that GH stimulated tyrosine phosphorylation of EGFR, whereas insulin did not (Fig. 1a, bottom). Moreover, immunodepletion of EGFR using specific antibody revealed that pp180 was almost entirely EGFR (Fig. 1b).

EGFR was tyrosine-phosphorylated in a time- and dose-dependent manner in response to GH in liver *in vivo* (Fig. 1c, d). In addition, EGFR was associated with Grb2 in response to GH (Fig. 1e). Shc was also tyrosine-phosphorylated in response to GH (Fig. 1f, lane 3). These results indicated that EGFR and Shc may represent two different pathways by which GH can activate MAP kinase (Fig. 1g), presumably through Grb2. Using specific antibodies, we found that GH could also induce tyrosine phosphorylation of IRS-1/IRS-2 in liver, albeit much less efficiently than insulin or the tyrosine

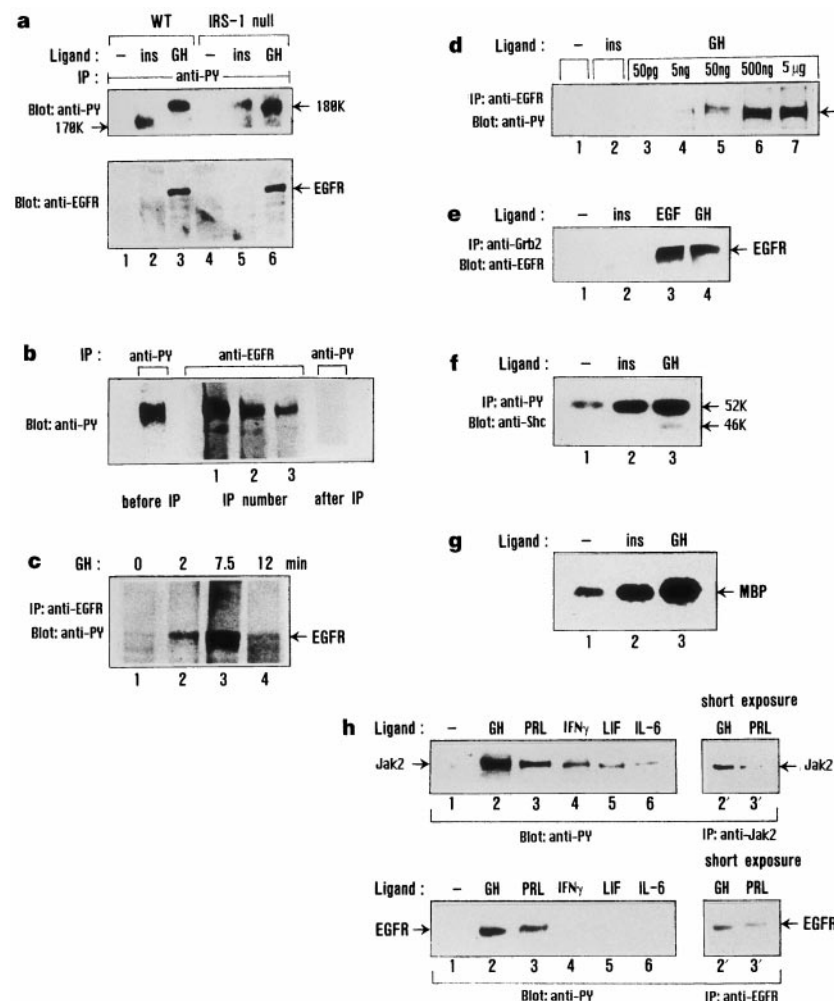


Figure 1 GH-stimulated tyrosine phosphorylation of EGFR and its association with Grb2, and concomitantly stimulated MAP kinase activity in liver *in vivo*. Wild-type (WT) or IRS-1-deficient mice (IRS-1 null) were treated with or without 1 U g⁻¹ body weight (BW) of insulin (ins) or 500 ng g⁻¹ BW of EGF for 5 min or 5 μ g g⁻¹ BW of hGH or PRL or 1 μ g g⁻¹ BW of IFN- γ or IL-6 or 2.5 μ g g⁻¹ BW of LIF for 7.5 min (**a**, **e**, **f**, **g**, **h**) or treated with 5 μ g g⁻¹ BW of hGH for the indicated time periods (**c**) or the indicated doses of hGH for 7.5 min (**d**). Liver lysates were immunoprecipitated (IP) with anti-PY (**a**, **f**) or anti-EGFR (**a**, bottom, **e**) or anti-Grb2 (**e**) or anti-Jak2 (**h**, top), followed by immunoblotting with anti-PY (**a**, top, **c**, **d**, **h**) or anti-EGFR (**a**, bottom, **e**) or anti-Shc (**f**). When bovine or human GH was injected, we obtained similar results (data not shown). **b**, Immunodepletion with anti-EGFR (see Methods). **g**, Liver extracts were immunoprecipitated with anti-MAP kinase (α C92; ref. 15), followed by kinase assay towards myelin basic protein (MBP). Labelled MBP was subjected to SDS-PAGE and autoradiography.

phosphorylation of EGFR stimulated with GH (unpublished observations).

We next tested whether other cytokine family members that activate Jak2 and have their own receptors in liver can stimulate tyrosine phosphorylation of EGFR. Interleukin-6 (IL-6), leukaemia inhibitory factor (LIF) and interferon- γ (IFN- γ) were able to activate Jak2 to 10–30% of the extent of GH and prolactin. However, only GH and prolactin, not IL-6, LIF or IFN- γ could stimulate tyrosine phosphorylation of EGFR (Fig. 1h), suggesting that it is specifically induced by GH and prolactin.

To study the role of tyrosine phosphorylation of EGFR in GH signalling, we used CHO cells stably transfected with the human GHR complementary DNA (CHO-GHR cells), which do not express endogenous EGFR. Wild-type (WT) EGFR or LacZ as a control was introduced into the cell line by infection with a recombinant adenovirus¹⁶. In CHO-GHR cells expressing LacZ, GH failed to induce tyrosine phosphorylation of EGFR (Fig. 2a, lane 3). In contrast, in CHO-GHR cells expressing WT EGFR, GH induced tyrosine phosphorylation of EGFR in a dose-dependent manner (Fig. 2a lane 6, and data not shown). Expression of WT EGFR in these cells was accompanied by GH-induced MAP kinase activation, significantly enhanced compared with cells expressing no EGFR at all the time points and GH concentrations studied (Fig. 2b, and data not shown). The association of tyrosine-phosphorylated proteins with Grb2 is a crucial step in the activation of the MAP-kinase cascade. Stimulation of CHO-GHR cells expressing LacZ by GH caused a 52K tyrosine-phosphorylated protein (Shc) to associate with a glutathione-S-transferase (GST)–Grb2 fusion protein (Fig. 2c, lane 2). In contrast, expression of WT EGFR resulted in GH-dependent association of EGFR with Grb2, which concomitantly reduced Shc binding to Grb2 (Fig. 2c,

lane 4), suggesting that there may be competition between EGFR and Shc. We next examined another cell line which expresses extremely low amounts of endogenous EGFR, INS-1 (ref. 17) (Fig. 2d, bottom). Stimulation of MAP kinase by GH was defective in INS-1 (Fig. 2e, lane 3), and could be restored by expression of EGFR (Fig. 2e, lane 6) through GH-induced tyrosine phosphorylation of EGFR (Fig. 2d, top, lane 6) and its association with Grb2 (data not shown).

To study the mechanisms of GH-induced tyrosine phosphorylation of EGFR, kinase-negative EGFR or an EGFR mutant with its main tyrosine-autophosphorylation sites substituted by phenylalanine (F5-EGFR) was introduced into CHO-GHR cells. In cells expressing kinase-negative EGFR, GH induced tyrosine phosphorylation of EGFR, which became associated with Grb2, and concomitantly stimulated MAP-kinase activity to an extent comparable to cells expressing WT EGFR (Fig. 3a, b, lanes 4, 6). In contrast, in CHO-GHR cells expressing F5-EGFR, there was little tyrosine phosphorylation of EGFR or association with Grb2 (Fig. 3a, lane 8), and GH stimulated MAP-kinase activity only to levels seen in cells expressing no EGFR (Fig. 3b, lanes 2, 8). These results indicate that full GH-induced activation of MAP kinase required tyrosine-phosphorylation sites on EGFR but did not depend on its intrinsic tyrosine-kinase activity. Tyrosine at residue 1,068 of EGFR is a major binding site for Grb2, when stimulated with EGF^{18,19}. A quadruple point mutant retaining one tyrosine at position 1,068 out of five tyrosine-autophosphorylation sites (F4/Y1068) was tyrosine-phosphorylated and could recruit Grb2 and induce activation of MAP kinase to almost the same extent as WT EGFR in response to GH in CHO-GHR cells (Fig. 3a, b, lanes 4, 10). This essential role of tyrosine-phosphorylation of EGFR was confirmed by the observation that overexpression of F5-EGFR (Fig. 3c, lanes 5, 6) significantly

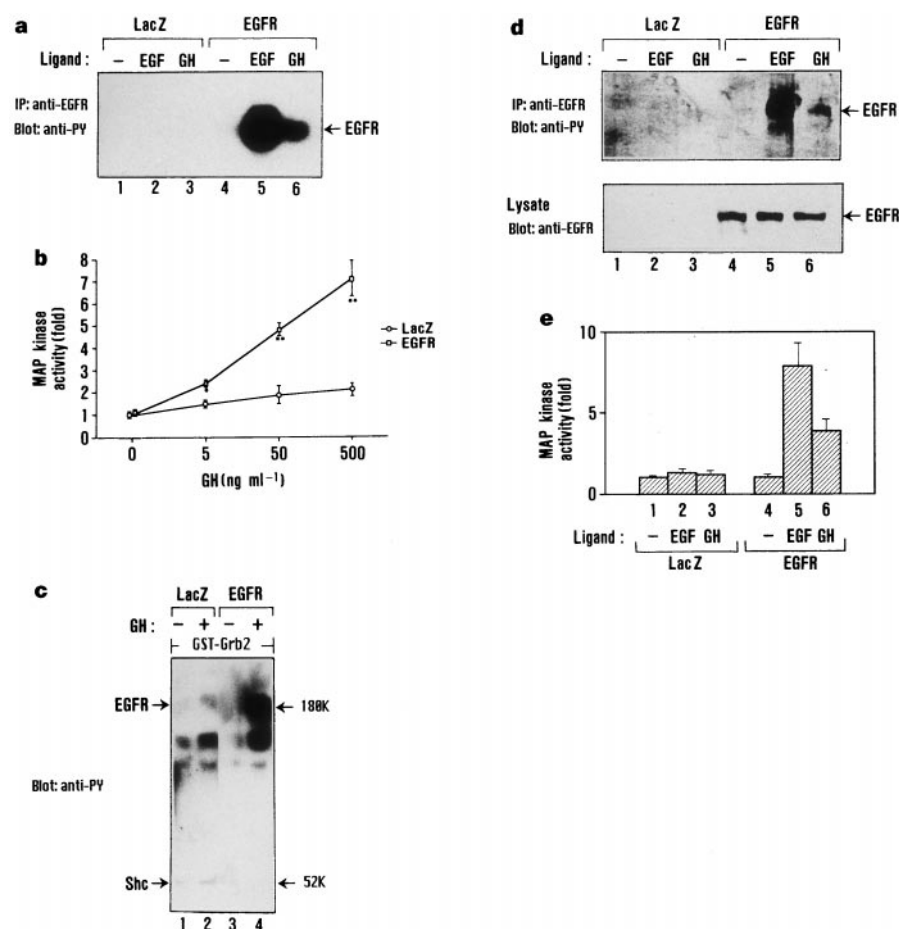


Figure 2 Tyrosine phosphorylation of EGFR enhances GH-induced MAP-kinase activation. **a, c, d**, Quiescent CHO-GHR cells (**a, c**) or INS-1(**d**) expressing wild-type EGFR or LacZ as a result of adenovirus-mediated gene transfer were stimulated for 5 min with EGF (100 ng ml⁻¹) or hGH (500 ng ml⁻¹). Upon lysis, precipitates with anti-EGFR (**a, d**, top) or GST-Grb2 (**c**) were immunoblotted with anti-PY (**a, c, d**, top). Equal expression was confirmed by probing with an anti-EGFR antibody in whole-cell lysates (**d**, bottom). **b, e**, Immune complex MAP-kinase assay. Quiescent CHO-GHR cells (**b**) or INS-1 (**e**) were stimulated with the indicated doses of hGH for 10 min (**b**) or with EGF (100 ng ml⁻¹) or hGH (500 ng ml⁻¹) for 10 min (**e**), and MAP kinase activity was determined as described¹⁵. Results are expressed as the ratio to the value of untreated CHO-GHR cells expressing LacZ. Each bar represents the mean $\pm$ s.e. of more than three independent experiments (* P < 0.01; ** P < 0.001; LacZ versus EGFR).

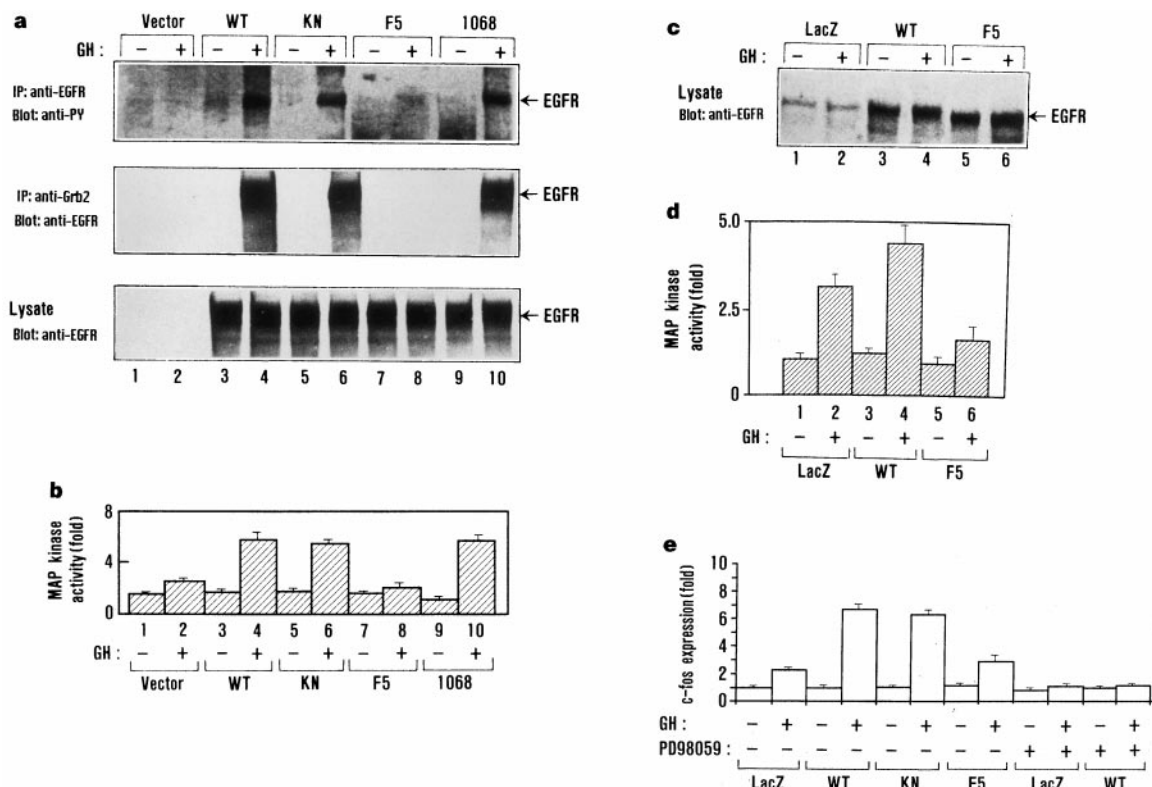


Figure 3 GH-induced full activation of MAP kinase and gene expression requires tyrosine-phosphorylation sites of EGFR but not EGFR intrinsic kinase activity. **a, b**, Quiescent CHO-GHR cells transfected with WT or EGFR mutant plasmids were stimulated with hGH (500 ng ml⁻¹) for 5 min (**a**) or for 10 min (**b**). Upon cell lysis, precipitated EGFR was immunoblotted with anti-PY (top), and immunoprecipitates by anti-Grb2 were immunoblotted with anti-EGFR (centre). Equal expression was confirmed by probing with anti-EGFR in whole-cell lysates (bottom) (**a**). MAP kinase activity was determined by immune-complex assay as described¹⁵. Results are expressed as the ratio to the value of untreated CHO-GHR cells expressing no EGFR. Representative data of one of three independent experi-

ments are shown; values are means $\pm$ s.d. (**b**). **c**, Expression of WT or F5-EGFR in 3T3 F442A cells via adenovirus-mediated gene transfer was confirmed by probing with anti-EGFR in whole-cell lysates. **d**, Immune complex MAP-kinase assay. Quiescent 3T3 F442A cells were stimulated for 10 min with hGH (500 ng ml⁻¹) and processed as in Fig. 2. **e**, GH-induced *c-fos* expression in CHO-GHR cells expressing LacZ or WT or kinase-negative (KN) or F5-EGFR with or without pretreatment with 30 μ M PD98059 for 60 min was determined as in Methods. Results are expressed as the ratio to the value of *c-fos* expression in untreated CHO-GHR cells expressing LacZ. Representative data from one of three independent experiments are shown; values are means $\pm$ s.d.

reduced GH-stimulated MAP-kinase activation in 3T3 F442A cells (Fig. 3d, lanes 2, 6), which express endogenous EGFR (Fig. 3c, lanes 1, 2).

To determine the importance of GH-induced tyrosine phosphorylation of EGFR in the function of GH, we studied GH-induced expression of *c-fos*. Full induction of *c-fos* by GH requires several regulatory elements in the promoter, including the *c-sis*-inducible element known as STAT (signal transducers and activators of transcription) 1 and 3-binding element and the ternary complex factor (TCF) element²⁰. Maximal transcriptional activity of both STAT proteins²¹ and TCF²² requires MAP kinases. As expected, in CHO-GHR cells expressing WT or kinase-negative EGFR, GH significantly stimulated *c-fos* expression compared with cells expressing LacZ or F5-EGFR, and this GH-induced increase in *c-fos* expression was almost abolished by pretreatment of cells with the MAP kinase kinase inhibitor PD98059 (Fig. 3e).

We next examined whether GH-induced tyrosine phosphorylation of EGFR is dependent on Jak2 by using γ 2A/GHR³ cells, which lack Jak2. GH was unable to induce tyrosine phosphorylation of EGFR (Fig. 4a, lane 3) nor to stimulate MAP-kinase activation of γ 2A/GHR cells transiently transfected with WT EGFR (data not shown), whereas complementation with a Jak2 expression plasmid restored these responses (Fig. 4a, lane 6, and data not shown). Furthermore, expression of Δ Jak2, which inhibits autophosphorylation of WT JAK2 in a dominant-negative manner²³, largely reduced GH-induced tyrosine phosphorylation of EGFR in CHO-

GHR transiently transfected with WT EGFR (Fig. 4b, lane 4). These results suggested that Jak2 kinase activity is required for tyrosine phosphorylation of EGFR in response to GH. G-protein-coupled receptors have been shown to mediate phosphorylation of EGFR through activation of Src tyrosine kinase²⁴, but GH did not stimulate the activity of Src kinase (data not shown). In addition, expression of Csk, which inactivates Src-family non-receptor tyrosine kinases by phosphorylating the regulatory carboxy-terminal tyrosine residue, reduced the Src kinase activity to only 20–30% of cells transfected with vector alone, whereas it had no apparent effect on GH-induced tyrosine phosphorylation of EGFR (data not shown). To clarify how GH induces phosphorylation of EGFR, we tested whether GH could induce the association of EGFR with GHR by co-immunoprecipitation experiments in COS cells transiently transfected with GHR and WT Jak2; we found that complex formation between GHR and EGFR was indeed induced in a GH-dependent manner (Fig. 4c, lane 2).

After GH binds to its receptor, the affinity of Jak2 for GHR appears to increase². To determine whether EGFR recruited to GHR by GH stimulation could be directly phosphorylated by Jak2, we carried out immune-complex kinase assays *in vitro*. The GST-cytoplasmic region of WT EGFR was phosphorylated by Jak2 immunisolated from CHO-GHR cell extracts treated with GH, and it was able to associate with Grb2 (Fig. 4d, lane 5). The amount of Jak2-induced phosphorylation of the GST-cytoplasmic region of the F4/Y1068-EGFR and its association with Grb2 were ~70–80%

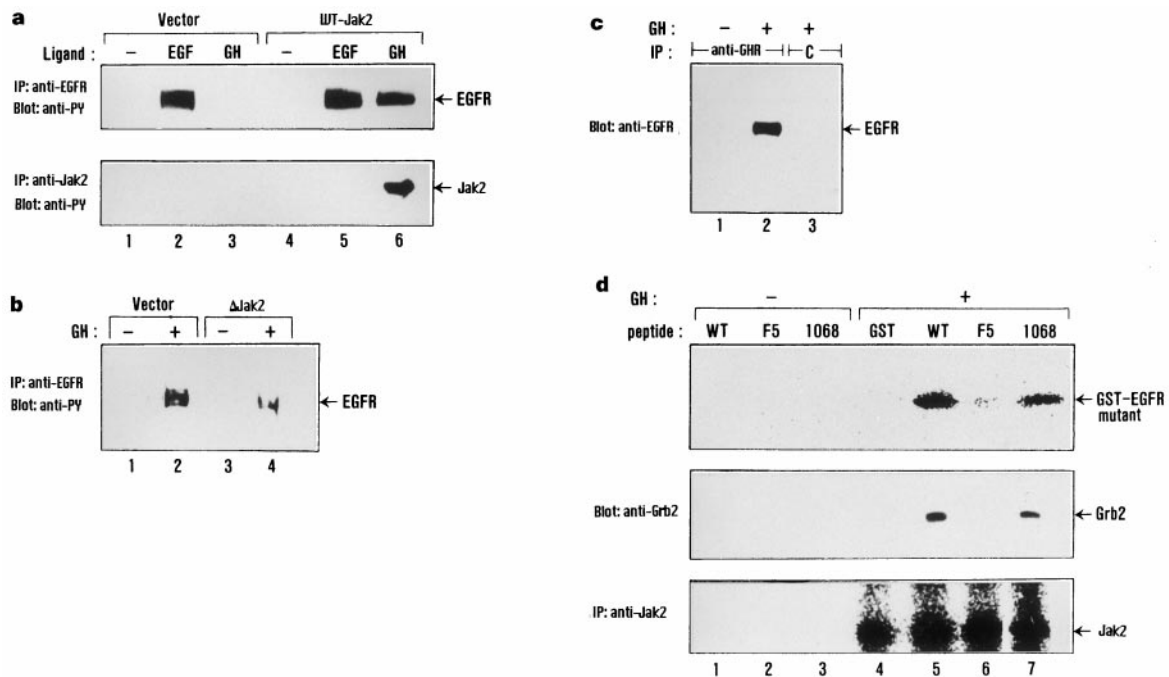


Figure 4 Tyrosine 1068 of EGFR was one of the major phosphorylation and Grb2-binding sites stimulated by GH via Jak2. **a–c**, WT EGFR plasmids were transfected into γ 2A/GHR with or without WT Jak2 plasmids (**a**) or CHO-GHR with or without Δ Jak2 plasmids (**b**). hGHR and WT Jak2 were cotransfected to COS cells (**c**). Quiescent cells were stimulated for 5 min with EGF (100 ng ml⁻¹) or hGH (500 ng ml⁻¹). Upon lysis, precipitates with anti-EGFR (**a**, top, **b**) or anti-Jak2 (**a**, bottom) or anti-GHR or control mouse IgG (**c**) were immunoblotted with

anti-PY (**a**, **b**) or anti-EGFR (**c**). **d**, *In vitro* Jak2 kinase assay. Top, phosphorylation of GST-EGFR-mutant fusion proteins induced by Jak2 immunoprecipitated from untreated or hGH-treated CHO-GHR cells. Centre, association of phosphorylated GST-EGFR-mutant fusion proteins with Grb2. Bottom, auto-phosphorylation of Jak2. Quiescent cells were stimulated with or without hGH and *in vitro* Jak2 kinase assays and *in vitro* association assays were done as described in Methods.

of that obtained for GST-WT-EGFR (Fig. 4d, lanes 5, 7), whereas these were almost undetectable for F5-EGFR (Fig. 4d, lane 6), which is consistent with *in vivo* data (Fig. 3a). These results indicated that Tyr 1068 may be one of the principal phosphorylation and Grb2-binding sites induced by Jak2 in response to GH, although other tyrosine phosphorylation sites may also be able to mediate association with Grb2 (refs 18, 19).

The EGF receptor is known to be important in signalling pathways triggered by ultraviolet irradiation²⁵ and by G-protein-coupled receptors²⁶, but it was considered that the intrinsic kinase activity of EGFR was necessary for signal transduction^{25,26}. More recently, it has been suggested that G-protein-coupled receptors may cause phosphorylation of EGFR by Src kinase, so independently of the intrinsic kinase activity of EGFR²⁴. Together with these observations, our findings provide the new paradigm that non-tyrosine-kinase receptors transduce signals through an associated non-receptor tyrosine kinase by phosphorylating the growth-factor receptor and utilizing it as a docking protein independent of its receptor tyrosine kinase activity. □

Methods

Materials. Murine IL-6, monoclonal anti-phosphotyrosine antibody (4G10), polyclonal antibody against mouse Jak2 and polyclonal antibody against human EGFR were all from Upstate Biotechnology; monoclonal antibody against EGFR (E 3138), bovine GH and ovine prolactin (PRL) were from Sigma; monoclonal antibody against EGFR (mab c11) was from Life Technologies; polyclonal antibody against EGFR was from Santa Cruz Biotechnology; monoclonal anti-hGHR antibody (mab5) was from Agen Biomedical (Brisbane, Australia); human GH (hGH) was a gift from Eli Lilly; murine IFN- γ , lipofectin and lipofectamine were from Gibco BRL; murine LIF was from Amrad (Melbourne, Australia); and PD98059 was from Calbiochem. All other materials were from the sources given in ref. 15. Mice were treated as described¹⁵.

EGFR mutants, adenovirus-mediated gene transfer and cell culture.

EGFRs^{19,28} were kindly provided by Y. Okabayashi and M. Kasuga. The recombinant adenoviruses Adex1CALacZ and Adex1CAEGFR were constructed by homologous recombination between the expression cosmid cassette and the parental virus genome¹⁶ and adenovirus-mediated gene transfer was done as described¹⁶. CHO cells were transfected with plasmids and selected as reported¹⁹ ($\sim 3.0 \times 10^4$ GH receptors per cell). When the adenovirus Adex1CALacZ was applied at 3×10^7 PFU cm⁻² per dish, LacZ was expressed in nearly 100% of CHO cells on post-infection days 1 and 2, and cells showed no significant changes, including in their morphology (data not shown). Thus, we applied recombinant adenoviruses at this dose and control cells were infected with Adex1CALacZ virus. INS-1 (ref. 17) and 3T3 F442A cells¹¹ were all grown as described.

MAP kinase assay, immunoprecipitation and immunoblotting. Immunoprecipitation, immunoblotting and MAP kinase assays have been described¹⁵. When immunoprecipitations were sequential, supernatants of liver homogenates from GH (5 μ g per g body weight)-stimulated mice were immunoprecipitated with anti-EGFR antibody. Immune complexes were collected with protein G-Sepharose; supernatants were reprecipitated with anti-EGFR antibody. Following three rounds of anti-EGFR immunoprecipitation, supernatants were immunoprecipitated with 4G10 and the immunoprecipitates from each step were immunoblotted with 4G10.

Expression of wild-type and kinase-inactive Jak2, hGHR and Csk. WT EGFR cDNA (3 μ g) in pRc/CMV was transfected into γ 2A/GHR⁵ or CHO-GHR cells with or without 3 μ g WT Jak2 plasmids or dominant-negative Jak2 (Δ Jak2; ref. 23), which lacks the C-terminal kinase domain under the SR α promoter, in 6-cm dishes by the lipofectamine method with modification²⁹ and the cells were stimulated as described⁵. The coding region of hGHR was isolated at *NheI* and *MluI* sites, and the insert was subcloned into pCI vector. COS cells were transfected with 3 μ g hGHR, 2 μ g WT Jak2 and 3 μ g Csk at 60–80% confluence in 3 ml of DMEM medium in 6-cm dishes by the calcium phosphate precipitation method; cells were stimulated as described³.

Generation of GST-EGFR fusion protein, *in vitro* Jak2 kinase assay, and

in vitro and in vivo association assays. The GST–cytoplasmic region of EGFR fusion proteins (amino acids 930–1,220, which do not include Lys 721, a key residue in the ATP-binding site) were generated as described²⁸. The *in vitro* Jak2 assay was done as before², with minimal modification. Proteins immobilized on anti-Jak2 antibody were mixed with 95 μ l of 0.1% Triton X-100 reaction buffer² containing 10 μ M ATP, 0.25 mCi ml⁻¹ [γ -³²P] ATP and 10 μ g GST–EGFR-mutant fusion protein. After reaction, the anti-Jak2 immune complexes were resolved by SDS–PAGE and the autophosphorylation of Jak2 was visualized by autoradiography. To collect GST–EGFR-mutant fusion proteins, the supernatant was incubated with 30 μ l glutathione–Sepharose beads; ³²P-labelled proteins were then resolved by SDS–PAGE and phosphorylation of GST–EGFR-mutant fusion proteins was visualized by autoradiography. When the *in vitro* association of phosphorylated GST–EGFR-mutant fusion proteins with Grb2 was determined, instead of being visualized by autoradiography, the fusion proteins were incubated with Grb2, processed as described¹⁹, and immunoblotted with anti-Grb2. To determine association of EGFR mutants with Grb2 *in vivo*, and the tyrosine-phosphorylation or MAP-kinase activity in cells expressing EGFR mutants, CHO–GHR cells were transfected with 6 μ g of pRc/CMV alone or with WT or EGFR mutant cDNA in pRc/CMV in 6-cm dishes by using the lipofectamine method with modification²⁹. Upon cell lysis, precipitated EGFR was immunoblotted with anti-PY (phosphotyrosine), and immunoprecipitates by anti-Grb2 were immunoblotted with anti-EGFR.

Luciferase assay. CHO–GHR cells were plated in 12-well dishes. One day after adenovirus-mediated gene transfer, cells were transfected by using the lipofectin procedure with 0.15 μ g fusion gene construct pFOS(-359)Luc³⁰. After incubation for 18 h with lipofectin in serum-free Ham's F-12 medium, transfected cells were pretreated for 60 min with or without 30 μ M PD98059, then treated with hGH (500 ng ml⁻¹) for 5.5 h. Luciferase was assayed as described²⁰.

Received 28 May; accepted 15 August 1997.

1. Ihle, J. N. *et al.* Signaling by the cytokine receptor superfamily: JAKs and STATs. *Trends Biochem. Sci.* **19**, 222–227 (1994).
2. Argentesinger, L. S. *et al.* Identification of JAK2 as a growth hormone receptor-associated tyrosine kinase. *Cell* **74**, 237–244 (1993).
3. Frank, S. J. *et al.* Regions of the JAK2 tyrosine kinase required for coupling to the growth hormone receptor. *J. Biol. Chem.* **270**, 14765–14785 (1995).
4. Winston, L. A. & Hunter, T. JAK2, Ras, Raf are required for activation of extracellular signal-regulated kinase/mitogen-activated protein kinase by growth hormone. *J. Biol. Chem.* **270**, 30837–30840 (1995).
5. Han, Y. *et al.* Participation of JAK and STAT proteins in growth hormone-inducing signaling. *J. Biol. Chem.* **271**, 5947–5952 (1996).
6. Ihle, J. N. STATs and MAPKs: obligate or opportunistic partners in signaling. *BioEssays* **18**, 95–98 (1996).
7. Burgering, B. M. T. & Bos, J. L. Regulation of Ras-mediated signalling: more than one way to skin a cat. *Trends Biochem. Sci.* **20**, 18–22 (1995).
8. van Biesen, T. *et al.* Receptor-tyrosine-kinase- and G β γ -mediated MAP kinase activation by a common signalling pathway. *Nature* **376**, 781–784 (1995).
9. Vanderkuur, J., Allevato, G., Billestrup, N., Norstedt, G. & Carter-Su, C. Growth hormone-promoted tyrosyl phosphorylation of Shc proteins and Shc association with Grb2. *J. Biol. Chem.* **270**, 7587–7593 (1995).
10. Ridderstrale, M., Degerman, E. & Tornqvist, H. Growth hormone stimulates the tyrosine phosphorylation of the insulin receptor substrate-1 and its association with phosphatidylinositol 3-kinase in primary adipocytes. *J. Biol. Chem.* **270**, 3471–3474 (1995).
11. Argentesinger, L. S. *et al.* Growth hormone, interferon- γ , and leukemia inhibitory factor promoted tyrosyl phosphorylation of insulin receptor substrate-1. *J. Biol. Chem.* **270**, 14685–14692 (1995).
12. Argentesinger, L. S., Norstedt, G., Billestrup, N., White, M. F. & Carter-Su, C. Growth hormone, interferon- γ , and leukemia inhibitory factor utilize insulin receptor substrate-2 in intracellular signaling. *J. Biol. Chem.* **271**, 29415–29421 (1996).
13. Tamemoto, H. *et al.* Insulin resistance and growth retardation in mice lacking insulin receptor substrate-1. *Nature* **372**, 182–186 (1994).
14. Araki, E. *et al.* Alternative pathway of insulin signalling in mice with targeted disruption of the IRS-1 gene. *Nature* **372**, 186–190 (1994).
15. Yamauchi, T. *et al.* Insulin signalling and insulin actions in the muscle and liver of insulin resistant IRS-1 deficient mice. *Mol. Cell. Biol.* **16**, 3074–3084 (1996).
16. Miyake, S. *et al.* Efficient generation of recombinant adenoviruses using adenovirus DNA-terminal protein complex and a cosmid bearing the full-length virus genome. *Proc. Natl Acad. Sci. USA* **93**, 1320–1324 (1996).
17. Sekine, N., Ullrich, S., Regazzi, R., Pralong, W. F. & Wollheim, C. B. Postreceptor signaling of growth hormone and prolactin and their effects in the differentiated insulin-secreting cell line, INS-1. *Endocrinology* **137**, 1841–1850 (1996).
18. Batzer, A. G., Rotin, D., Urena, J. M., Skolnik, E. Y. & Schlessinger, J. Hierarchy of binding sites for Grb2 and Shc on the epidermal growth factor receptor. *Mol. Cell. Biol.* **14**, 5192–5201 (1994).
19. Okutani, T. *et al.* Grb2/Ash binds directly to tyrosine 1068 and 1086 and indirectly to tyrosine 1148 of activated human epidermal growth factor receptors in intact cells. *J. Biol. Chem.* **269**, 31310–31314 (1994).
20. Chen, C., Clarkson, R. W. E., Xie, Y., Hume, D. A. & Waters, J. M. Growth hormone and colony-stimulating factor-1 share multiple response elements in the *c-fos* promoter. *Endocrinology* **136**, 4505–4516 (1995).
21. Wen, Z., Zhong, Z. & Darnell, J. E. Jr Maximal activation of transcription by Stat1 and Stat3 requires both tyrosine and serine phosphorylation. *Cell* **82**, 241–250 (1995).

22. Gille, H., Sharrocks, A. D. & Shaw, P. E. Phosphorylation of transcription factor p62TCF by MAP kinase stimulates ternary complex formation at *c-fos* promoter. *Nature* **358**, 414–417 (1992).
23. Watanabe, S., Itoh, T. & Arai, K. JAK2 is essential for activation of *c-fos* and *c-myc* promoters and cell proliferation through the human granulocyte-macrophage colony-stimulating factor receptor in BA/F3 cells. *J. Biol. Chem.* **271**, 12681–12686 (1996).
24. Luttrell, L. M., Rocca, G. J. D., Biesen, T. V., Luttrell, D. K. & Lefkowitz, R. G. β γ subunits mediate src-dependent phosphorylation of the epidermal growth factor receptor. *J. Biol. Chem.* **272**, 4637–4644 (1997).
25. Sachsenmaier, C. *et al.* Involvement of growth factor receptors in the mammalian UV response. *Cell* **78**, 963–972 (1994).
26. Daub, H., Weiss, F. U., Wallasch, C. & Ullrich, A. Role of transactivation of the EGF receptor in signalling by G-protein-coupled receptors. *Nature* **369**, 557–560 (1996).
27. Honegger, A. M. *et al.* Point mutation at the ATP binding site of EGF receptor abolishes protein-tyrosine kinase activity and alters cellular routing. *Cell* **51**, 199–209 (1987).
28. Okabayashi, Y. *et al.* Tyrosines 1148 and 1173 of activated human epidermal growth factor receptors are binding sites of Shc in intact cells. *J. Biol. Chem.* **269**, 18674–18678 (1994).
29. Meunier-Durmort, C., Ferry, N., Hainque, B., Delattre, J. & Forest, C. Efficient transfer of regulated genes in adipocytes and hepatoma cells by the combination of liposomes and replication-deficient adenovirus. *Eur. J. Biochem.* **237**, 660–667 (1996).
30. Tanaka, T. *et al.* Evi-1 raises AP-1 activity and stimulates *c-fos* promoter transactivation with dependence on the second zinc finger domain. *J. Biol. Chem.* **269**, 24020–24026 (1994).

Acknowledgements. We thank I. Kerr and G. Stark for γ 2A/GHR, C. Wollheim for INS-1, H. Green for 3T3 F442A cells, J. Ihle for Jak2 cDNA, K. Arai for expression vectors of Jak2 and dominant-negative Jak2 (Δ Jak2), H. Sabe for Csk expression vector, I. Saitoh for AdexiCALacZ and the expression cosmid cassette, Y. Okabayashi and M. Kasuga and S. Kakinuma for technical assistance. This work was supported by grants from the research fellowships of the Japan Society for the Promotion of Science for Young Scientists (to T.Y.), from the Ministry of Education, Science, Sports, and Culture and the Ministry of Health and Welfare of Japan (to T.K.) and for diabetes research from the Taisho Pharmaceutical Co. Ltd. (to T.K.).

Correspondence and requests for materials should be addressed to T.K. (e-mail: kadowaki-3im@h.u-tokyo.ac.jp).

Peptide bond formation by *in vitro* selected ribozymes

Biliang Zhang & Thomas R. Cech

Howard Hughes Medical Institute, Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309-0215, USA

An attractive solution to the problem of the origin of protein synthesis in an evolving 'RNA world' involves catalysis by nucleic acid without assistance from proteins^{1,2}. Indeed, even the modern ribosome has been considered to be fundamentally an RNA machine³, and the large ribosomal subunit can carry out peptidyl transfer in the absence of most of its protein subunits⁴. Successive cycles of *in vitro* selection and amplification^{5–7} have been used to find RNAs that perform many biochemical reactions^{8–16}, including transfer of an RNA-linked amino acid to their own 5'-amino-modified terminus¹⁵. Here we demonstrate the *in vitro* selection of ribozymes (196 nucleotides) that perform the same peptidyl transferase reaction as the ribosome: that is, they can join amino acids by a peptide bond. Like ribosome substrates, one amino acid (N-blocked methionine) is esterified to the 3'(2')-O of adenosine, whereas the acceptor amino acid (phenylalanine) has a free amino group. Our best characterized ribozyme recognizes the amino-acid ester substrate by binding its adenosine moiety, and is therefore capable of utilizing Leu- and Phe- as well as Met-derived substrates. Such lack of specificity with respect to the amino acid is a feature necessary for a generalized protein-synthesizing enzyme.

We synthesized a pool of 1.3×10^{15} different RNA molecules containing two regions of randomized sequence (70 and 72 nucleotides) flanked by constant regions. The template for *in vitro* transcription was a synthetic DNA library constructed by using the polymerase chain reaction (PCR) (see Methods). Transcription was performed in the presence of guanosine-5'-monophosphorothioate (GMPS) to introduce a 5'-phosphorothioate. The 196-mer 5'-GMPS-RNA pool was then reacted with *N*-bromoacetyl-*N*'-phenylalanyl-cystamine to convert to 5'-Phe-SS-RNA. Thus, each RNA had a Phe at its 5' end that provided a free amino group capable of functioning as an aminoacyl acceptor (Fig. 1).

To select active RNA, the 5'-Phe-SS-RNA pool was incubated

with the substrate AMP-Met-Bio (2'(3')-N-biotinylamidocaproyl-L-methionyl-adenosine-5'-monophosphate) in the presence of Mg^{2+} at 25 °C. AMP-Met-Bio is a simple analogue of the natural initiator formyl methionine(fMet)-transfer RNA (Fig. 1). RNA able to form a peptide bond became covalently linked to a biotin tag and was selected by binding to a streptavidin column. The active RNA molecules were eluted from the streptavidin column by reduction of the disulphide linkage with dithiothreitol (DTT). Selected RNAs were reverse-transcribed to complementary DNA, amplified by PCR, transcribed to 5'-GMPS-RNA, and finally converted to 5'-Phe-SS-RNA to enter the next cycle of selection. At the seventh generation, the enriched RNA pool showed a band-shift upon gel electrophoresis that represented a streptavidin-RNA complex (Fig. 2a). The retarded band disappeared after treatment with DTT (Fig. 2a, lane 9c'), which indicated that the biotin was linked to the RNA on the external (Phe side) of the disulphide linker. This suggested that the expected peptide bond had formed between the carbonyl carbon of the amino acid of the substrate (AMP-Met-Bio) and the amino group (phenylalanine) of 5'-Phe-SS-RNA. The active RNA reached 10.4% of the ninth generation RNA pool. After 19 cycles of selection, the activity of the pool RNA had increased markedly, being 1.7×10^4 -fold more active than the seventh-generation RNA based on initial rate measurements.

The nineteenth-generation cDNA molecules were cloned. PCR-amplified DNAs from 75 plasmids were used as templates for run-

off transcription by T7 RNA polymerase. Kinetic analysis of the individual RNA species revealed that 9 of 75 had better activity than the final RNA pool, while 18 clones were inactive (Fig. 2b). Sequence analysis (24 clones) showed that the most active RNAs comprised at least two sequence classes. The sequence of one of these classes, which included clones 8, 14, 25, 45 and 60, is given as Supplementary information. The deletion of ~20 nucleotides from either the 5' or the 3' end, or from both ends of clone-25 RNA decreased its catalytic activity dramatically (data not shown). Therefore, both the 5' and 3' constant regions of this ribozyme are necessary for formation of the active structure.

Kinetic analysis with clone-25 RNA and various concentrations of AMP-Met-Bio gave a typical Michaelis-Menten curve (Fig. 3a). The resulting kinetic parameters were: $k_{cat} = 0.05 \text{ min}^{-1}$ and $K_m = 190 \mu\text{M}$. Compared with the uncatalysed reaction of N-acetylaminoacyl-tRNA transfer with glycine ethyl ester¹⁷, the rate enhancement of the peptide bond formation catalysed by clone-25 RNA is $\sim 1 \times 10^6$. The reaction of clone-25 RNA required Mg^{2+} and reached a maximal rate when $[Mg^{2+}] \geq 100 \text{ mM}$. As with other ribozymes¹⁸, Mg^{2+} may be required for proper RNA folding and may also be involved directly in the chemical step of peptide bond formation. Divalent metal ions also are required for peptidyl transfer to puromycin by the ribosome¹⁹.

To test whether the chemical identity of the linker affected the reaction, it was shortened from 13 to 5 interatomic bonds (Fig. 1 legend). This linker modification of the clone-25 ribozyme abolished activity. Furthermore, the free, linker phenylalanine added *in trans* inhibited the reaction of the clone-25 ribozyme (50% inhibition between 5 and 10 mM), but no inhibition was observed by free phenylalanine or phenylalanine amide. One likely explanation is that the linker provides specific contacts with the RNA that are required for proper positioning of the 5' amino group of 5'-Phe-SS-RNA in the active site of the ribozyme. Because inhibition by the free linker

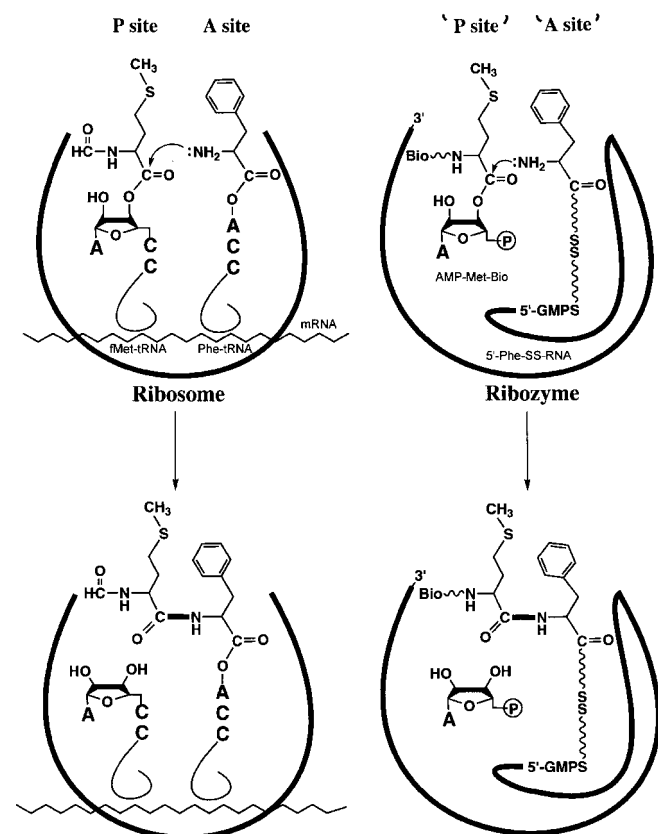


Figure 1 Peptide bond formation by a ribosome (left) and by a ribozyme (right). The process on the ribosome involves two binding sites, the P (peptidyl) and A (aminoacyl) sites. In the ribozyme reaction, 5'-Phe-SS-RNA (functioning like aminoacyl-tRNA in the A site of the ribosome) provides the amino group that can attack the aminoacyl carbonyl carbon of AMP-Met-Bio (mimicking the fMet-tRNA or peptidyl-tRNA in the P site of the ribosome) to form a peptide bond (thick line). The standard linker (wavy vertical line) joining GMPS and Phe was $\text{CH}_2\text{C}(\text{O})\text{NHCH}_2\text{CH}_2\text{SSCH}_2\text{CH}_2\text{NH}$. One experiment used a shortened linker, $\text{SCH}_2\text{CH}_2\text{NH}$, attached directly to the 5' G rather than through the thiophosphate (GMPS).

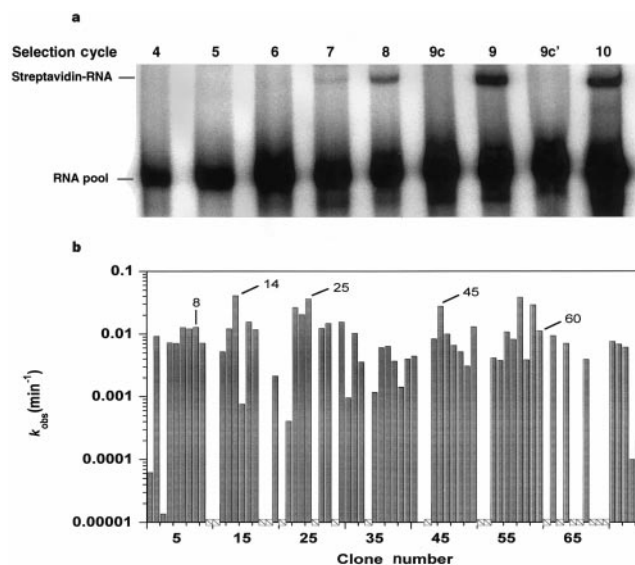


Figure 2 Activity of the selected RNA pool and transcripts from individual clones from generation 19. **a**, Polyacrylamide gel showing the shift of a portion of the radiolabelled RNA after incubation with AMP-Met-Bio and then with streptavidin. Selection reactions were done at 25 °C with $4 \mu\text{M}$ 5'-Phe-SS-RNA and 8 mM AMP-Met-Bio substrate for 20 h, except for selection cycle 10, which was done with $1 \mu\text{M}$ 5'-Phe-SS-RNA and 1 mM AMP-Met-Bio for 5 h. Control reactions for pool-9 RNA: lane 9c, no streptavidin; lane 9c', RNA treated with DTT before streptavidin incubation. **b**, Distribution of the observed rate constants of 75 cloned RNAs. 5'-Phe-SS-RNA ($4 \mu\text{M}$), prepared from PCR-amplified DNA template by *in vitro* transcription, was incubated with $400 \mu\text{M}$ AMP-Met-Bio substrate in reaction buffer at 25 °C. RNAs with no detectable activity ($<10^{-6} \text{ min}^{-1}$) are indicated by squares with diagonal line.

requires it to compete with the linker attached intramolecularly to the ribozyme, the interaction energy may be much greater than would be indicated by an apparent K_d of 5–10 mM. In addition, the length of the linker may be critical for allowing the amino group of Phe to reach the active site. Thus, instead of the reactive groups being aligned by local RNA–RNA base pairing, as in many previous ribozyme reactions^{8,9,11,13–15,20,21}, interactions with the linker may provide alignment.

The peptidyl transferase reaction catalysed by clone-25 RNA was inhibited by adenosine-5'-monophosphate (AMP), but not by other nucleoside-5'-monophosphates. A K_i of about 188 μ M for AMP was measured (data not shown). The equivalence of K_m for AMP-Met-Bio and K_i for AMP suggests that most of the substrate binding energy comes from interactions with the AMP, not the amino acid or biotin. The ability of RNA to form specific mononucleotide binding sites is well known^{22,23}. Methionine and *N*-formyl-methionine at 1 mM concentration did not inhibit the peptidyl transferase reaction, although higher concentrations of methionine may be required to reveal potential weak interactions between the amino acid of the substrate and the ribozyme.

Several control assays validated the ribozyme-mediated formation of a peptide bond. When 5'-GMPS-RNA (without linked phenylalanine) was incubated with the AMP-Met-Bio substrate under the 5'-Phe-SS-RNA reaction conditions, no product was detected (Fig. 3b, lane 10). Also no product formed without Mg^{2+} , or at zero time of incubation (Fig. 3b, lanes 2 and 7). The disappearance of the product upon treatment with DTT suggested that the biotin group was transferred to the 5'-Phe-SS-RNA externally to the disulphide bond (Fig. 3b, lane 8), consistent with formation of the desired peptide bond. To confirm the formation of the Met-Phe peptide bond, a large-scale reaction was performed to obtain enough material for HPLC and mass spectroscopy. After being released from the RNA by DTT treatment, the RNA-catalysed reaction product had a retention time (23 min) by high-performance liquid chromatography (HPLC) and a mass peak (695 $[M + H^+]$) by mass spectroscopy (MS) identical to those of the chemically synthesized authentic compound (Fig. 4).

Several 2'(3')-*N*-biotinylamidocaproyl-L-aminoacyl-AMP substrates (Leu, Phe, Met and Lys) were synthesized to evaluate the specificity of the peptide-bond-forming activity of clone-25 RNA. The four substrates were all active, albeit with different reaction rates (Fig. 5). Met and Leu were the best peptidyl donors, and kinetic measurements showed that they reacted with the same rate constant but to different final extents. Lys and Phe were less efficient donors for peptidyl transfer. Although the ribosome catalyses over 20 types of amino-acid transfer reaction during normal protein synthesis, the donor activity of various amino-acid-charged substrates is dependent on the identity of the amino-acid side chain when the tRNA is replaced with smaller RNA fragments^{24–26}. In these experiments, the derivatives of acetylated (Ac) Phe were poor donors, and the highest donor activity was observed with derivatives of fMet and AcLeu²⁴, the same amino-acid preference as for the clone-25 ribozyme reported here.

The peptidyl transferase reaction of the selected ribozymes is fundamentally similar to that carried out by the ribosome (Fig. 1). The 3'(2')-aminoacyl ester of free adenosine bound by the selected RNA is analogous to the aminoacyl adenosine at the end of tRNA bound in the P site (peptidyl site) of the ribosome. In contrast to the ribosome, the 'A-site' (aminoacyl site) amino acid of the ribozyme is tethered by a non-nucleotide linker. However, we found that the linker provides proper spacing or alignment of the acceptor amino acid (Phe), which appears to be analogous to the role of the A-site tRNA in the ribosome. Although the selected ribozymes contain only ~6% of the number of nucleotides of the contemporary large subunit ribosomal RNA, they also perform a simpler reaction. There is no codon–anticodon recognition on our ribozyme, nor is there translocation. Furthermore, the reaction rates are slow compared to

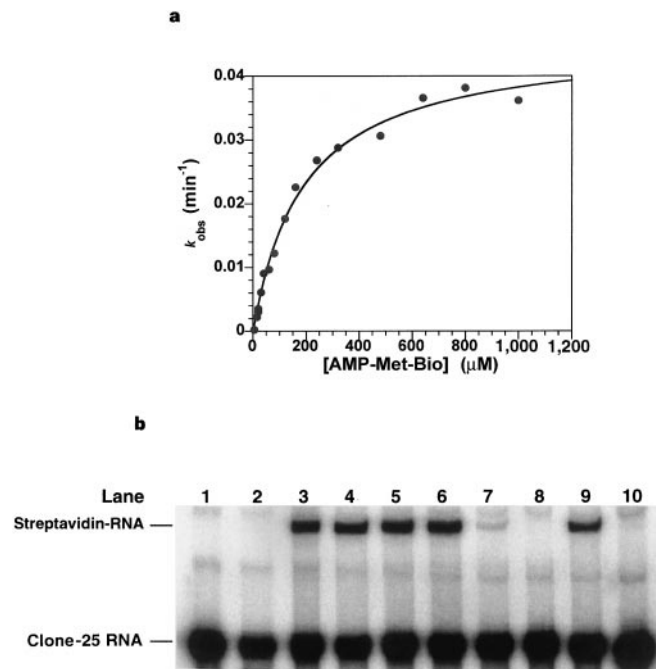


Figure 3 Formation of biotinylated dipeptide catalysed by clone-25 RNA. **a**, Reaction kinetics of clone-25 RNA at 25 °C. Reactions were run in 4 μ M 5'-Phe-SS-RNA and 50 mM Mg^{2+} at pH 7.4. Best-fit values of K_m and k_{cat} were obtained by fitting the data to the Michaelis–Menten equation by KaleidoGraph program. **b**, The standard reaction contained 0.4 mM AMP-Met-Bio substrate and 4 μ M clone 25 5'-Phe-SS-RNA in 300 mM KCl and 200 mM Mg^{2+} . After 2.5 h at 25 °C, aliquots were removed from the reaction mixture for streptavidin gel-shift assays. Alterations in the standard conditions were: lane 1, no substrate; lane 2, no Mg^{2+} ; lane 3, 30 mM Mg^{2+} ; lane 4, 60 mM Mg^{2+} ; lane 5, 100 mM Mg^{2+} ; lane 6, standard reaction; lane 7, zero time; lane 8, 3 μ mol DTT; lane 9, 0.1 μ mol DTT; and lane 10, 4 μ M 5'-GMPS-RNA (no 5'-Phe).

the 15–20 peptide bonds formed per second on a ribosome¹⁷. Nevertheless, it is now clear that even in a limited portion of RNA sequence space, there are many RNA sequences that possess peptidyl transferase activity.

The ribosome may achieve peptidyl transfer simply through precise orientation of substrates, or it may also provide chemical groups to stabilize high-energy oxyanionic intermediates and transition states¹⁷. Study of the mechanism of peptidyl transfer by *in vitro* selected ribozymes may help reveal the mechanistic features of ribosome-catalysed protein synthesis. □

Methods

Oligonucleotide synthesis. The oligodeoxynucleotides PM1, 5'-AGCGAA TTCTAATACGACTCACTATAGGGAGAGCCATACCTGAC-3'; PM2, 5'-CACGGATCCTGACGACTGAC-3'; N70DNA, 5'-GGGAGAGCCATACCTGAC-N₇₀-CAGGTTACGCATCC-3'; and N72DNA, 5'-CACGGATCCTGACGACTGAC-N₇₂-GGATGCGTAACCTG-3' were synthesized on an Applied Biosystems DNA synthesizer, model 394. A mixture of phosphoramidites in a ratio 3:3:2:2 (A:C:G:T) was used to synthesize the random positions (N).

Aminoacyl substrate synthesis. AMP-Met-Bio was synthesized by esterification of AMP with *N*-biotinylamidocaproyl-L-methionine²⁷. The other AMP-aminoacyl-Bio compounds were similarly synthesized from the corresponding *N*-biotinylamidocaproyl-L-amino acids. The products were purified on a silica gel column eluted with *n*-butanol:acetic acid:water ratio of 78:5:17 (v/v). The purity of products was confirmed by ¹H-NMR and electrospray ionization/MS.

Initial RNA pool. The full-length 222-mer double-stranded (ds) DNA was synthesized by PCR amplification with 120 μ g N70DNA and 120 μ g N72DNA as templates. PCR with primers PM1 and PM2 was run in a total volume of

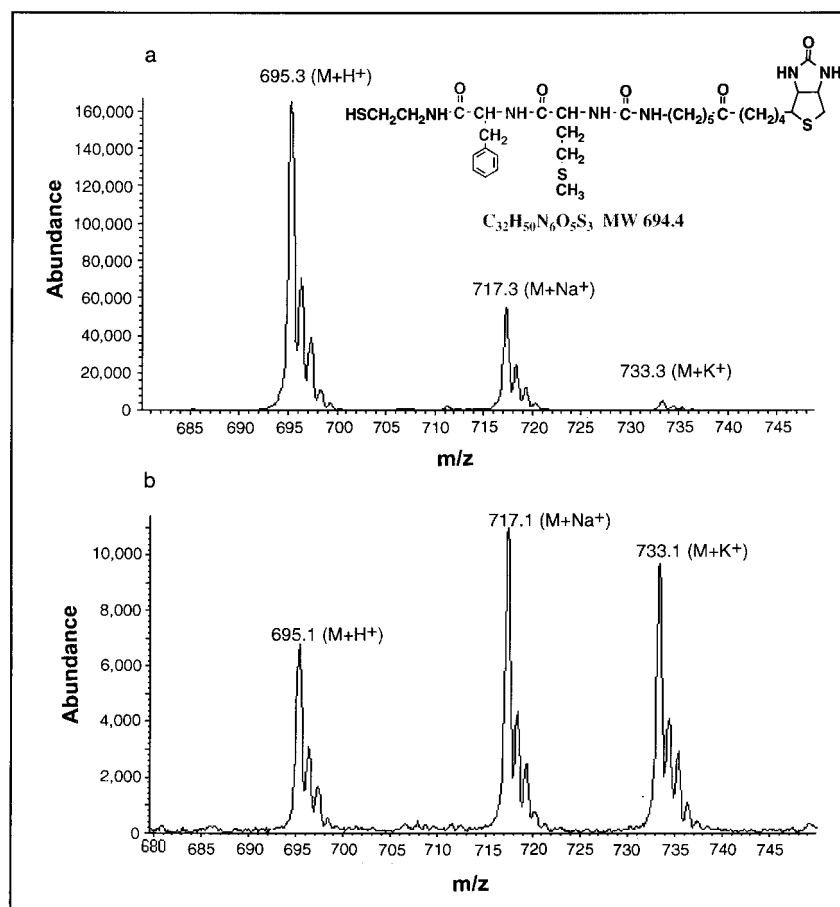


Figure 4 Validation of the peptide bond formed by catalysis of clone-25 RNA by HPLC-ESI/MS analysis of the dipeptide product. **a**, Authentic compound *N*-biotinylamidocaproyl-L-methioninyl-L-phenylalanine-2'-pyridyldithioethyl amide (Bio-Met-Phe-SS-Py) was synthesized by coupling of *N*-biotinylamidocaproyl-methioninyl-*N*-hydroxyl-succinimide ester with *L*-phenylalanine-2'-pyridyl-dithioethyl amide (made by reacting *t*-Boc-*L*-phenylalanine-*N*-hydroxyl succinimide ester with pyridyldithioethylamine). This product was then reacted with DTT to give the compound shown. **b**, 5'-*N*-biotinylamidocaproyl-L-methioninyl-L-phenylalaninyl-SS-GMPS-RNA (product of the clone-25 RNA reaction) was also reacted with DTT. The HPLC separation was performed on a Hewlett-Packard 1090 liquid-chromatography system with a C18 reverse-phase column. The elution gradient was 100% water (1 min), 56.7% acetonitrile (35 min) and 100% acetonitrile (41 min). Mass spectrometry was done on a Hewlett-Packard 5989B mass spectrometer. Both the HPLC and MS were controlled by the HP Chemstation software.

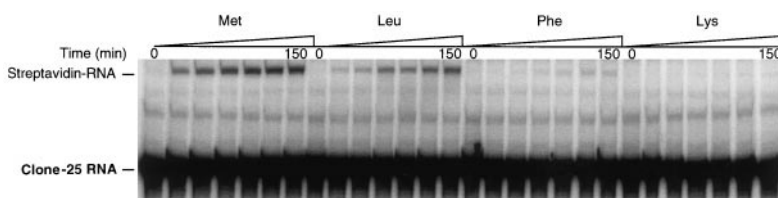


Figure 5 Amino-acid specificity of peptide bond formation catalysed by clone-25 ribozyme. The time course of reactions with 8 μ M clone-25 RNA and 500 μ M 3'(2')-*N*-biotinylamidocaproyl-aminoacyl-AMP substrates (Met, Leu, Phe and Lys) in 300 mM KCl, 100 mM MgCl₂ and 50 mM HEPES (pH 7.4) at 25 °C.

39.2 ml in 196 tubes, each containing 200 μ l of 10 mM Tris-HCl, pH 8.3; 50 mM KCl; 0.01% (w/v) gelatin; 0.05% Tween 20; 0.2 mM dNTPs; 1.0 μ M primers and 3.0 mM MgCl₂, in 10 cycles (94 °C, 1 min; 52 °C, 1 min; 72 °C, 2 min). The PCR product was extracted once with chloroform:phenol (1:1) and once with chloroform:isoamyl alcohol (24:1), precipitated with 3.0 volumes of ethanol, resuspended in water, and used for the next step without further purification. The initial RNA pool was prepared from 800 μ g PCR-amplified DNA by *in vitro* transcription in a 20-ml reaction mixture (40 mM Tris-HCl, pH 7.5, 10 mM DTT, 4 mM spermidine, 0.05% Triton X-100, 12 mM MgCl₂, 1 mM each ATP, CTP, UTP, 8 mM GMPS, 0.4 mM GTP, 100 μ Ci [α -³²P]ATP and T7 RNA polymerase (supplied by A. Gooding)) for 19 h at 37 °C (for the short-linker variant, GMPS was replaced by 5'-deoxy-5'-thioguanosine-5'-monophosphate). 800 U of RNase-free DNase I (Promega) and CaCl₂ (final concentration, 10 mM) were added to digest DNA templates. After 2 h, 0.4 ml 0.5 M EDTA (pH 8.0) and 0.6 ml 5.0 M NaCl were added, and the RNA was precipitated with 3.0 volumes of ethanol. The 5'-GMPS-RNA was purified by electrophoresis on a denaturing 8% polyacrylamide gel with 80–90% yield. The 5'-GMPS-RNA (50 μ M) was reacted with *N*-bromoacetyl-*N*'-phenylalanyl-cystamine (10 mM) in chemical linking buffer (20 mM HEPES, pH 7.7, 150 mM NaCl, 1 mM EDTA) for 1 h at room temperature and overnight at 4 °C, giving a yield of 40–60%. After phenol extraction, the 5'-Phe-SS-RNA pellet was collected by precipitation with ethanol and used for the selection.

Selection. In the first 9 cycles, 4 μ M 5'-Phe-SS-RNA was incubated with 8 mM

AMP-Met-Bio substrate in reaction buffer (50 mM HEPES, pH 7.4, 300 mM KCl, 50 mM MgCl₂) for 20 h at 25 °C. The reaction mixture was extracted with chloroform:phenol (1:1) and chloroform:isoamyl alcohol (24:1), and precipitated with 3.0 volumes of ethanol. The pellet was redissolved in 1.0 ml binding buffer (1.0 M NaCl, 20 mM HEPES, pH 7.4, 5 mM EDTA) and transferred to a tube containing 1 ml (5 ml for first cycle selection) of 50% streptavidin-agarose gel slurry (prewashed twice with binding buffer). The slurry mixture was mixed gently for 40–60 min by a roller, loaded on a column, and then washed with 20 ml binding buffer, 20 ml water, 20 ml 7.0 M urea, 20 ml water. Bound RNA was eluted with 3.0 ml 100 mM DTT in binding buffer. The eluted RNA was precipitated with 100 μ g glycogen and 2.5 volumes of ethanol. The eluted RNA was converted to cDNA by reverse transcription. A 20- μ l annealing mixture (40 mM Tris-HCl, pH 8.3, 8.0 mM DTT, 48 mM NaCl, 5 pmol 3'-primer) was heated for 90 s at 90 °C and slowly cooled to room temperature. Then 50 μ l of reverse transcription (RT) mixture (20 μ l annealing mix, 0.4 mM dNTPs, 40 mM Tris-HCl, pH 8.3, 8.0 mM DTT, 48 mM NaCl, 2.4 mM MgCl₂, 4 U of avian myeloblastosis virus reverse transcriptase) was incubated for 30 min at 40 °C. The cDNA was amplified by PCR in a 100- μ l reaction (50 μ l RT reaction mix: 2.4 mM MgCl₂, 23 mM Tris-HCl, pH 8.3, 12.5 mM KCl, 0.2 mM dNTPs, 1.0 μ M primers, 5 U *Taq* DNA polymerase) and purified as described, transcribed, and linked with phenylalanine for the next selection. In subsequent selection cycles, 1–2 μ M 5'-Phe-SS-RNA and 200–400 μ M AMP-Met-Bio were incubated for 1–5 h; other conditions were the same as those already described.

Kinetic studies. All reactions were carried out with uniformly [α - 32 P]ATP-labelled ribozyme and excess AMP-aminoacyl-Bio substrate (≥ 100 -fold) in 50 mM HEPES (pH 7.4), 300 mM KCl and 50 mM MgCl₂ at 25 °C. The RNA was preincubated with buffer and Mg²⁺ for 10 min at 50 °C, and cooled to room temperature (~20 min). The reaction was initiated by addition of AMP-aminoacyl-Bio substrate. At each time point, 7 or 8 aliquots of 2 μ l were removed from 20 μ l reaction mixtures and quenched in dry ice with one volume of 'stop' buffer containing 100 mM HEPES (pH 7.4) and 100 mM EDTA. Samples were incubated with 10 μ g streptavidin in binding buffer at room temperature before mixing with 0.25 volumes of 90% formamide loading buffer. Biotinylated RNA products were separated by electrophoresis on 8% polyacrylamide/7 M urea gels (1,000 V at 4 °C for 2 h). The fraction of product formation relative to total substrate and product at each time point was quantified with a Molecular Dynamics PhosphorImager. Reactions reached 30–50% completion and rates were corrected for this end point.

Received 6 August; accepted 17 September 1997.

1. Woese, C. R. in *The Genetic Code: The Molecular Basis for Genetic Expression* 179–195 (Harper & Row, New York, 1967).
2. Crick, F. H. C. The origin of the genetic code. *J. Mol. Biol.* **38**, 367–379 (1968).
3. Noller, H. F. in *The RNA World* (eds Gesteland, R. F. & Atkins, J. E.) 137–156 (Cold Spring Harbor Laboratory Press, Plainview, NY, 1993).
4. Noller, H. F., Hoffarth, V. & Zimmick, L. Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* **256**, 1416–1419 (1992).
5. Ellington, A. D. & Szostak, J. W. *In vitro* selection of RNA molecules that bind specific ligands. *Nature* **346**, 818–822 (1990).
6. Robertson, D. L. & Joyce, G. F. Selection *in vitro* of an RNA enzyme that specifically cleaves single-stranded DNA. *Nature* **344**, 467–468 (1990).
7. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505–510 (1990).
8. Bartel, D. P. & Szostak, J. W. Isolation of new ribozymes from a large pool of random sequences. *Science* **261**, 1411–1418 (1993).
9. Lorsch, J. R. & Szostak, J. W. *In vitro* evolution of new ribozymes with polynucleotide kinase activity. *Nature* **371**, 31–36 (1994).
10. Prudent, J. R., Uno, T. & Schultz, P. G. Expanding the scope of RNA catalysis. *Science* **264**, 1924–1927 (1994).
11. Dai, X.-C., de Mesmaeker, A. & Joyce, G. F. Cleavage of an amide bond by a ribozyme. *Science* **267**, 237–240 (1995).
12. Illangasekare, M., Sanchez, G., Nickles, T. & Yarus, M. Aminoacyl-RNA synthesis catalyzed by an RNA. *Science* **267**, 643–647 (1995).
13. Wilson, C. & Szostak, J. W. *In vitro* evolution of a self-alkylating ribozyme. *Nature* **374**, 777–782 (1995).
14. Eklund, E. H. & Bartel, D. P. RNA-catalysed RNA polymerization using nucleoside triphosphates. *Nature* **382**, 373–376 (1996).
15. Lohse, P. A. & Szostak, J. W. Ribozyme-catalysed amino-acid transfer reactions. *Nature* **381**, 442–444 (1996).
16. Wecker, M., Smith, D. & Gold, L. *In vitro* selection of a novel catalytic RNA: characterization of a sulfur alkylation reaction and interaction with a small peptide. *RNA* **2**, 982–994 (1996).
17. Nierhaus, K. H., Schultz, H. & Cooperman, B. S. Molecular mechanisms of the ribosomal peptidyltransferase center. *Biochem. Int.* **1**, 185–192 (1980).
18. Pyle, A. M. Ribozymes: a distinct class of metalloenzymes. *Science* **261**, 709–714 (1993).
19. Maden, B. E. H. & Monro, R. E. Ribosome-catalyzed peptidyl transfer: effects of cations and pH value. *Eur. J. Biochem.* **6**, 309–316 (1968).
20. Piccirilli, J. A., McConnell, T. S., Zaug, A. J., Noller, H. F. & Cech, T. R. Aminoacyl esterase activity of the *Tetrahymena* ribozyme. *Science* **256**, 1420–1424 (1992).
21. Cech, T. R. in *The RNA World* (eds Gesteland, R. F. & Atkins, J. E.) 239–269 (Cold Spring Harbor Laboratory Press, Plainview, NY, 1993).
22. Bass, B. L. & Cech, T. R. Specific interaction between the self-splicing RNA of *Tetrahymena* and its guanosine substrate: implications for biological catalysis by RNA. *Nature* **308**, 820–826 (1984).
23. Sassanfar, M. & Szostak, J. W. An RNA motif that binds ATP. *Nature* **364**, 550–553 (1993).
24. Monro, R. E., Cerna, J. & Marker, K. A. Ribosome-catalyzed peptidyl transfer: substrate specificity at the P-site. *Proc. Natl Acad. Sci. USA* **61**, 1042–1049 (1968).
25. Mercer, T. F. B. & Symons, R. H. Peptidyl-donor substrates for ribosomal peptidyl transferase: chemical synthesis and biological activity of *N*-acetyl aminoacyl di- and trinucleotides. *Eur. J. Biochem.* **28**, 38–45 (1972).
26. Krayevsky, A. A., Kukhanova, M. K. & Gottikh, B. P. Peptidyl transferase center of bacterial ribosomes: substrate specificity and binding site. *Nucleic Acids Res.* **2**, 2223–2236 (1975).
27. Azharyev, A. V. *et al.* Synthesis of amino acid esters of nucleoside 5'-phosphates and their investigation with PMR and CD spectra. *Nucleic Acids Res.* **4**, 2223–2234 (1977).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from Mary Sheehan at the London editorial office of Nature.

Acknowledgements. We thank K. Goodrich and E. Podell for oligonucleotide synthesis; R. Barkley and O. Averin for HPLC-MS spectra; and B. Golden, R. Gottlieb, G. Joyce, S. Seiwert and O. Uhlenbeck for comments on the manuscript. B.Z. is supported by a postdoctoral fellowship from the National Institute of General Medical Science, NIH. T.R.C. is an investigator of the Howard Hughes Medical Institute and an American Cancer Society Professor.

Correspondence and requests for materials should be addressed to T.R.C. (e-mail: thomas.cech@colorado.edu).

correction

Synthesis of epothilones A and B in solid and solution phase

K. C. Nicolaou, N. Winssinger, J. Pastor, S. Ninkovic, F. Sarabia, Y. He, D. Vourloumis, Z. Yang, T. Li, P. Giannakakou & E. Hamel

Nature **387**, 268–272 (1997)

We wish to add that, as well as a total synthesis of epothilone B, reference 19 includes biological data for compound **23** and other congeners similar to those reported in our Letter. □

retraction

High levels of genetic change in rodents of Chernobyl

Robert J. Baker, Ronald A. Van Den Bussche, Amanda J. Wright, Lara E. Wiggins, Meredith J. Hamilton, Erin P. Reat, Michael H. Smith, Michael D. Lomakin & Ronald K. Chesser

Nature **380**, 707–708 (1996)

We previously reported an elevated DNA substitution rate in the mitochondrial cytochrome *b* gene of voles from Chernobyl. Our conclusion was based on a higher level of variation in rodents living in the zone of radioactivity than in rodents from control regions. Those DNA sequences were obtained by manual sequencing of cloned PCR products. We have subsequently attempted to replicate these findings with direct, automated sequencing of PCR products from the same animals. We found discrepancies between these sequences generated by direct automated methods and those previously reported. Re-examination of the original autoradiograms revealed that the elevated levels of variation were largely a product of human error in scoring, aligning and recording the sequences. This error has been confirmed by automated sequencing of a selected sample of the original clones on which the manual sequences were based. There remains a higher rate of substitution in the nine animals from the radioactive zone than in the ten animals from the control population, but the difference is not statistically significant with this sample size. There is variation among the clones from the mice living in the radioactive regions, including clones from the embryos. The two mutations reported in the embryos do not qualify as substitutions based on the standard set forth in the methods and materials that changes be present in multiple clones. The significance (if any) of this variation remains to be determined. The data do not support a mutation rate as high as that reported, and we retract that conclusion. □